COMMUNICATION CONCEPTS 2

# Information

L. David Ritchie

# Information

# COMMUNICATION CONCEPTS

This series reviews enduring concepts that have guided scholarly inquiry in communication, including their intellectual evolution and their uses in current research. Each book is designed to provide organized background reading for those who intend further study of the subject.

EDITOR

Steven H. Chaffee, *Stanford University*

ASSOCIATE EDITORS

Charles R. Berger, *University of California, Davis*
Joseph N. Cappella, *University of Pennsylvania*
Robert P. Hawkins, *University of Wisconsin-Madison*
Mark R. Levy, *University of Maryland, College Park*
Neil M. Malamuth, *University of California, Los Angeles*
Jack McLeod, *University of Wisconsin-Madison*
Peter Monge, *University of Southern California*
Byron Reeves, *Stanford University*
Michael Schudson, *University of California, San Diego*
Ellen Wartella, *University of Illinois at Urbana-Champaign*

EDITORIAL ASSISTANT

Caroline Schooler, *Stanford University*

COMMUNICATION CONCEPTS 2

# Information

L. David Ritchie

SAGE PUBLICATIONS
*The International Professional Publishers*
Newbury Park London New Delhi

Copyright © 1991 by Sage Publications, Inc.

All rights reserved. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

*For information address:*

SAGE Publications, Inc.
2455 Teller Road
Newbury Park, California 91320

SAGE Publications Ltd.
6 Bonhill Street
London EC2A 4PU
United Kingdom

SAGE Publications India Pvt. Ltd.
M-32 Market
Greater Kailash I
New Delhi 110 048 India

Printed in the United States of America

ISBN 0-8039-3904-3 (c) ISBN 0-8039-3905-1 (p)

ISSN 1057-7440

**FIRST PRINTING, 1991**

Sage Production Editor: Astrid Virding

---

Citation Instructions:

When citing a **Communications Concepts** issue, please follow this reference style:

Ritchie, L. David (1991). *Communication Concepts 2: Information.* Newbury Park, CA: Sage.

# *Contents*

Foreword
*Steven H. Chaffee and Joseph Cappella* vii

Preface viii

1. Information in Communication Science 1
   *Concepts* 1
   *Information in Recent Communication Literature* 3
   *The Mathematical Theory of Communication* 5
   *Historical Perspective* 6
   *Signal Transmission as Metaphor* 8
   *Matching Concept to Theory* 9

2. Communication: Signal Transmission and Interpretation 10
   *Communication* 10
   *Signal Transmission* 11
   *Interpretation* 16
   *Data, Relevance, and Information* 18

3. Measuring the Variety of a Set 20
   *The Power of a Code* 21
   *A Basic Metric: The Binary Set* 22
   *Codes with Unequal Distribution of Elements* 27
   *What H Measures* 30
   *Some Useful Characteristics of H* 32

4. Redundancy: How Structure Affects Variety 32
   *Code Efficiency* 33
   *Redundancy* 33
   *Structure* 35
   *How Internal Structure Contributes to Communication* 36
   *How Internal Structure Affects H* 38

5. Structure and Relevance 42
*The Structure of the Transmission System* 43
*The Structure of a Message* 47
*Social Structure* 47

6. Uncertainty 50
*Defining Uncertainty* 50
*Uncertainty and False Paradoxes* 53
*Managing Uncertainty in Human Communication* 55
*Uncertainty in Communication: Causes and Remedies* 57
*Measuring Uncertainty* 60

7. Summary: Distinctions and Connections 61
*Distinctions* 62
*Connections* 65

References 68

Author Index 73

About the Author 75

# *Foreword*

Each volume in the **Communication Concepts** series deals at length with an idea of enduring importance to the study of human communication. Through analysis and interpretation of the scholarly literature, specialists in each area explore the uses to which a major concept has been put, and point to promising directions for future work.

Information is clearly a—perhaps *the*—central concept in the study of communication. We asked L. David Ritchie to range across several levels of conceptual usage in this one small book: information's technical meaning in engineering; the complex meanings of information and various pseudonyms—*uncertainty, structure, entropy, redundancy*—in specialized academic studies; and its metaphorical usage by communication theorists. He has, with our encouragement, concentrated largely on the middle of this list, developing the academic and theoretical meanings that are most applicable to the human side of communication. His examples cover a wide range, including such domains as interpersonal communication, group dynamics, mass media, social structure, social influence, decision making, organizational communication, and the history of technology. This book offers a synthesis that is relevant to the interests of virtually every student of communication processes.

Far from having an agreed-upon, self-evident meaning, *information* is a challenging topic for conceptualization that has occupied many of the best minds in the field of communication for half a century. Professor Ritchie deserves our thanks for summarizing and resolving many difficult issues with this book. This volume should help focus the thinking of the next generation of scholars, who, we are certain, will find further stimulation in the concept of information.

—Steven H. Chaffee, *Series Editor*
Joseph Cappella, *Associate Editor*

# *Preface*

The ideas presented in this monograph began to take shape while I was a graduate student at Stanford University; the first drafts of several chapters were written at the University of Wisconsin—Madison. I am deeply indebted, for my general approach to concept explication and theory development, to the inspiration and teaching of Steven H. Chaffee, Donald F. Roberts, and Richard F. Carter. Both Peter Monge and Joseph Cappella were kind enough to read the manuscript in considerable detail and spare me thereby from at least some of my most embarrassing mistakes. The ideas presented herein have also benefited from extensive discussions with many colleagues and friends, including Klaus Krippendorff, Vincent Price, John Peters, David Mortensen, James Dillard, Isabelle Bauman, and Eugene Buder. My wife, LaJean Humphries, has been a consistent source of advice and encouragement. I hope the product will prove to be worthy of their trust and friendship.

# INFORMATION

L. DAVID RITCHIE

## *1. Information in Communication Science*

What is information? We routinely say that one book is more informative than another, that one medium is a better source of information, that a meeting was uninformative. We get our living in an information age by doing information work. We buy, sell, conceal, and reveal information, and we seek information to reduce our uncertainty. Information is central to an understanding of human communication.

Communication scientists and scholars sometimes use the word information in a broad way, relying on the general sense of the word to convey their meaning. A student or reader may be expected to understand ideas such as an "informative speech," "information resources," and "information gatherers" without the need of a precise conceptual definition. However, when information is itself at the center of a theoretical discussion (e.g., how much information can be conveyed by one speaking style versus another or by one medium versus another; how does information affect behavior), the concept needs to be more precisely defined.

### Concepts

Every time we describe something, we use concepts: Common examples include color, size, time, and emotions. Any concept is

an abstraction from the detail of everyday perception, a way of isolating certain features that are common to many different experiences. The concepts we choose reflect our implicit theories about what is important: If we choose to describe something by its color, we are expressing an implicit theory that color tells us something important about that thing. Whenever experiences are classified according to some quality, and that quality is taken as the basis for distinguishing between concepts (for example, when we distinguish "love" from "lust"), the act of classifying and naming makes a claim that is unavoidably theoretical, a claim that the concepts defined by a certain quality can help us understand, predict, and control events.

Use of a scientific concept makes a claim of a similar sort. A concept is defined by a theory and unavoidably carries embedded within it the assumptions of that theory. The theory implicitly or explicitly tells us what to notice and what to ignore. A scientific concept must express something that a theory says is important, something that a theory tells how to detect and interpret.

Information is sometimes understood as a concept appropriated from electronic engineering, along with the assumptions and perspectives of signal transmission theory. To use information thus, as a concept defined by signal transmission theory, is to accept the implicit claim that, at some empirically observable level, human communication is "nothing but" signal encoding, transmission, and decoding, and to take on the burden of explaining human communication in mechanistic terms. It is doubtful that many researchers would willingly assume either the burdens or the constraints implied by defining information in such terms: The alternative is to define information in terms of human communication theory.

If information is to be useful as a concept in the scientific study of human communication, it must be distinguished both from the irrelevant or unsupported assumptions conveyed by the way scientists in other disciplines use the word and from the welter of meanings given the word in everyday conversation. The concept of information needs to be distinguished from related concepts such as data, uncertainty, and structure as well as from the statistical tools that have been developed to aid in the observation of these concepts. A primary purpose of this volume is to explain these distinctions and why they are important.

Other purposes are to clear away misconceptions from the everyday use of the word as well as from casual discussions in popular writings and even in some textbooks, and to show how the concept of information can connect various aspects of the communication process in a coherent way.

## Information in Recent Communication Literature

Introductory speech communication textbooks routinely include sections on "informative speaking" and discuss information gathering as part of preparing a speech. Informative speeches explain, define, clarify, or instruct (DeVito, 1988; Hybels & Weaver, 1986; Seiler, 1988). Information to be used in preparing a speech includes facts, evidence, data, quotations, examples, and statistics (Busby & Majors, 1987). Interpersonal communication textbooks discuss information primarily in terms of how individuals learn about one another. Information includes spoken and written words, distance, facial expressions, and bodily movements (Busby & Majors, 1987) as well as emotions and intentions (Tubbs & Moss, 1983).

Information is also described in terms of answering questions and reducing uncertainty. Informational speeches answer specific questions; informational interviews answer questions and reveal things (Pearson & Nelson, 1988). Being informed "helps to reduce our uncertainty of things we know little about" (Seiler, 1988, p. 243). To answer questions and reduce uncertainty, information must connect the unknown with what is already known (Hybels & Weaver, 1986). The informational speech must be relevant: It must originate in the speaker's experience and connect with the audience's experience (Pearson & Nelson, 1988). Mass communication textbooks discuss information as facts and news (DeFleur & Dennis, 1988), as ideas and influences that shape social behavior, and as a means of resolving ambiguity and reducing threat (DeFleur & Ball-Rokeach, 1989).

Recent articles in communication journals reveal a similar pattern. Information includes data, knowledge, and opinions (Douglas, 1985) that answer questions (Baxter & Wilmot, 1984; Slater, 1990) and reduce uncertainty (Baxter & Wilmot, 1984; Douglas, 1985; Hewes, Graham, Monsour, & Doelger, 1989; Kellermann,

1986). Some authors identify information with detail (Abbott, 1989; Kurke, Weick, & Ravlin, 1989). Others distinguish information from data and emphasize the importance of relevance or connectedness to a context or to other information (Gandy, 1989; Griffith, 1989).

Information is frequently understood by reference to "information processing," drawing on the digital computer as a metaphor for cognitive and social processes. Information is a means to order experience and to plan and control action (Beniger, 1986; Couch, 1988). Information serves bureaucratic surveillance purposes (Gandy, 1989) and reinforces social stratification (Braman, 1989; Murdock & Golding, 1989; Schiller, 1989).

In research and teaching, communication scientists tend to identify information as those aspects of messages that tell something, answer questions, inform someone about something. Information is always relevant to the context of some human activity, a context that may be quite personal at times, quite social at others.

In technical writing, information is identified with the statistical distribution of elements in a set and consequently with the power of a code to identify certain things or ideas. Human communication is often explained through metaphors based on electronic data processing and signal transmission devices. The computer and signal transmission metaphors are explicit in many introductory textbooks (e.g., Bittner, 1988; Brooks & Heath, 1989; DeFleur & Dennis, 1988) and are prominently featured in several textbooks on communication theory (e.g., Borman, 1989; Littlejohn, 1989; Severin & Tankard, 1988). These writers explain the concept of information in terms of statistical uncertainty or variety and explicitly link information to Shannon's (1948, 1949) "mathematical theory of communication." In doing so, they are pursuing an approach first suggested by Weaver (1949) in his introductory interpretation of Shannon's ideas, enthusiastically endorsed by Schramm (1955), and developed in considerable detail by Cherry (1957/1982).

About a third of this volume will be devoted to Shannon's ideas, a fact that might seem anomalous, given Shannon's (1949, p. 31) repeated avowal that his theories had nothing to do with meaning and were concerned only with the problem of "reproducing at one point either exactly or approximately a message

selected at another point." Whether he willed it or not, Shannon's ideas have proven extremely influential among scholars and researchers interested in human communication, and the concept of information cannot be adequately addressed without discussing Shannon and the subsequent uses (and misuses) of his ideas.

## The Mathematical Theory of Communication

Shannon's (1948, 1949) mathematical theory of communication was directed toward the problem of determining the maximum capacity of a channel (such as a telephone or telegraph circuit) for transmitting messages. He was particularly interested in the relationship among three factors: the way messages are encoded, the presence of noise (any condition that randomly alters part of the signal), and the capacity of a channel. The meaning of the messages is totally irrelevant; all that matters is the statistical description of the typical messages produced by the code.

To estimate the efficiency of a code based on a set of N signs (such as the 26 letters of the alphabet), Shannon employed a statistical formula for entropy, first applied to signal transmission by Hartley (1928): $H = -\Sigma p(i) \log_2 p(i)$, where i = 1 through N and p(i) refers to the probability that the $i^{th}$ sign will appear in any position in a typical message. If H is calculated for the 26 letters of the alphabet, for example, p(t) is the probability that any letter selected at random from a typical message will be "t." In general, *entropy* refers to the degree of randomness or dispersion among elements of some set: The letters "a" through "z" are elements of the alphabet, and the achievable motions (position and velocity) of every atom in a closed container are elements of a set of thermodynamic states. H increases as the number of elements increases and, for any given number of elements, H has its maximum value if each element is equally likely to appear in any position. Organization and structure constrain the order in which elements may appear and hence make some elements more probable in certain positions. For example, English spelling requires that each word has at least one vowel. Consequently, H can also be considered a measure of disorganization: The more organized a system, the lower the value of H. In thermodynamics, the

amount of disorganization in a physical system, as measured by H, is referred to as the *entropy* of the system.

Most codes constrain the order in which elements may appear (e.g., in English, "q" is almost always followed by "u," and "h" is likely to follow "c," "s," or "t"). These constraints result from the rules that constitute the structure of the code. Structure reduces the amount of variety that will be observed and results in redundancy. The more redundancy in a code, the more elements it will require to encode a typical message from another code. On the other hand, if some elements are lost or randomly altered by noise in the transmission channel, the internal structure can be used to identify and correct the errors. (These ideas are more formally developed in Chapters 2 through 5.)

Shannon (1948) used these ideas to demonstrate a series of theorems about the relationship of the statistical characteristics of a code and the amount of random disturbance in a transmission channel to the rate at which messages can be transmitted through the channel. Shannon's theorems have played an important role in electronics and computer science, but they have not proven particularly useful in the study of human communication.

## Historical Perspective

In Gardner's (1987) account, Shannon made two important contributions to the scientific understanding of how systems represent, manipulate, and transmit information. The first was his master's thesis at MIT (Shannon, 1938), in which he showed how relationships in abstract logic can be expressed by series of "binary" switches. (A binary switch has two values or states, such as on or off, open or closed, true or false.) As a practical consequence of Shannon's work, any process that can be described in logical form can be performed by a digital computer.

Shannon's second contribution was to derive a formula for quantifying the minimum number of binary switches (e.g., true or false, open or closed, 1 or 0) required to identify an element of any set with a known distribution and to apply that formula to the solution of some crucial problems in cryptography and electronic circuit design (Shannon, 1948). When entropy is expressed in logarithms of base 2, $H = -\Sigma p(i) \log_2 p(i)$, H expresses the number

of binary digits, or bits, required to identify an average element of the set. For example, in a deck of playing cards, only one bit is required to identify a card as red or black. Two bits would suffice to identify the suit: one to identify the suit as red or black, and a second to identify a red suit as hearts or diamonds or a black suit as spades or clubs.

Shannon's second essay was widely read by mathematicians, philosophers, psychologists, and others who incorporated its key ideas into disciplines as diverse as electronic engineering, economics, biology, psychology, and the new discipline of cognitive science. H came to be regarded as a measure of information, and the discipline built around the application of H to diverse scientific problems came to be known as information theory.

*"Information theory" in human communication science.* The extension of "information theory" (i.e., signal transmission theory) into the social sciences was facilitated when the University of Illinois Press reprinted Shannon's (1948) landmark paper in a slender volume along with an essay by Weaver (Shannon & Weaver, 1949). Weaver acknowledged that Shannon's theorems were intended to apply only to the technical processes of encoding and signal transmission but insisted that it would be easy to extend them to the questions of meaning and effectiveness that interest communication researchers.

The idea that the precision of engineering could be applied to social problems had long appealed to philosophers and social scientists (e.g., Dewey, 1921/1948); now physicists and engineers were beginning to think along similar lines (see Wiener, 1950). Weaver's arguments fit their program perfectly. If, as Weaver insisted, Shannon's restrictive assumptions could be ignored and his theorems extended beyond the technical problems of signal transmission, these theorems might provide the basis for advances in "social engineering" that would duplicate and perhaps surpass the spectacular wartime and postwar achievements of electronic engineering.

*Research in human communication.* Weaver's ideas helped precipitate a flurry of activity as communication researchers attempted to apply Shannon's ideas to human communication. Most of the research used H as a measure of variety; only a few tried to connect

variety with the ideas of meaning and understanding. Smith (1966) published a collection of chapters inspired by information theory that became a popular textbook for communication theory courses, but the chapters that discussed information presented no empirical research, and the empirical chapters said little about information.

Hsia (1968a, 1968b, 1968c, 1969) framed the comparison of auditory with visual media as a question of neurological transmission capacity, but the results are explainable without signal transmission theory (Jester, 1968). Garner (1962) made a more extensive effort along similar lines but admitted that he was unable to see what the concept of information added to his work. The greatest successes have been achieved in projects in which H does not necessarily have anything to do with information. Chaffee and Wilson (1977) used H to measure the diversity of political agendas in a community. Watt and his colleagues used H to measure variety and complexity in television programming (Krull, Watt, & Lichty, 1977; Watt & Krull, 1974; Watt & Welch, 1983), and Finn (1986) used H to measure unpredictability in news articles. Shannon's insights about redundancy were successfully applied to the problem of measuring readability (Darnell, 1970; Dickens & Williams, 1964; Lowry & Marr, 1975; Lynch, 1974; Taylor, 1953) and writing style (Paisley, 1966).

Signal transmission provides a tempting and evocative metaphor (Finn & Roberts, 1984), but signal transmission is only one part of the process of human communication. The use of signal encoding, transmission, and decoding as a model for all of human communication obscures the part-whole relationship and probably does more to hinder than to advance communication science. At the very least, more extensive discussion of information and wider use of H in communication research have been discouraged by the continuing expectation that information and H should be somehow connected to each other, as they are in signal transmission theory.

## Signal Transmission as Metaphor

As the theoretical basis of electronic engineering took shape over the first half of the century (see Pierce & Noll, 1990), key concepts (including communication and information) were given names borrowed from the social realm, based on the root metaphor

"signal transmission is communication." The same root metaphor suggested labels for other phenomena encountered by electronic engineers: "noise," "uncertainty," "feedback," and "redundancy," to mention only a few.

The machine had long been a popular metaphor for human beings, as had mechanical signaling for human communication. There was little protest when the metaphor "signal transmission is communication" was inverted into "communication is signal transmission": An electronic metaphor for communication was entirely consistent with the emerging metaphor "the brain is a computer." The perceived similarities between electronics and human communication led to a double metaphor, in which words like *noise, uncertainty, redundancy,* and *information* were applied as metaphors to the human realm from which they had originally been borrowed, with definitions and assumptions imported from the technical realm.

Metaphors highlight certain aspects of experience, hide other aspects, and thereby shape thought (Lakoff & Johnson, 1980). The effect is particularly pronounced when the origins of language are forgotten and metaphors are treated as a simple expression of "the way things are." Lakoff and Johnson give the example "argument is war." So effectively does this metaphor highlight the combative, competitive aspect of argument and hide the cooperative, investigative aspect that we sometimes find it nearly impossible to think of argument as anything but a contest in which there can be only one winner—and one loser. Similarly, "communication is signal transmission" highlights the mechanistic aspects of communication and hides the interpretive and social aspects. "Argument" and "war" are both instances of a broader concept—disagreement—but "signal transmission" is part of the broader concept "communication." Using part of the process as a metaphor for the whole makes it doubly difficult to understand the complete process: How can the role of signal transmission in communication be understood when communication is itself described as signal transmission?

## Matching Concept to Theory

For the concept of information to be useful in theories of human communication, its definition needs to express the attributes of a

message that the theory predicts will be important. Information needs to be defined in terms of the full process by which humans communicate, not restricted to the limited subprocess of signal transmission. Information needs to be distinguished from H as a measure of variety and transmission system capacity. Once the distinction is clearly drawn between H and information, communication researchers can feel free to use H without explaining how their work is related to information and to define information according to the theoretical model they mean to test, without feeling constrained to explain what their idea of information has to do with H.

# 2. *Communication: Signal Transmission and Interpretation*

The word *communication* comes from the root *common* and has to do with making something (e.g., information) common. After something is told, it is known to both the person who tells and the person who is told. The root word *common* also leads to the word *community:* The process of telling and making information common implies some sort of relationship between teller and told. The study of communication is a study of the processes by which information is told and community is maintained.

## Communication

The process of telling something requires that a message be created in such a form that it can be perceived and interpreted. The message itself can take many forms, as long as it is perceivable as a message. Ordinarily we think of transmission as the reproduction of a message at some remote place, but in some sense every act of communication includes transmission. In human speech, a pattern of sound waves is transmitted

through the air. In writing, visible patterns are transmitted, if only to a later time.

For the telling to be completed, some other person must perceive and interpret the message. Perception includes recognition: The perceiver must have some basis for discriminating between the events and patterns that are messages and those that are not messages. Interpretation often includes or is preceded by decoding, a process of translating from one sort of pattern to another. For example, the ear decodes speech from a pattern of sound waves to a pattern of nerve impulses. Interpretation also includes finding or making some relationship between the perceived message and other events, actions, and ideas.

Communicating creates relationships between what is perceived or known by one person and what is perceived or known by another; it also relies on pre-existing relationships. The receiver and originator of a message must work from some common understanding of what sorts of patterns are used to communicate and how these patterns are related to other events. Communication has to do with community both in the sense that it relies on having something in common in the first place and in the sense that it can influence what the communicants subsequently have in common.

All of these processes are apparent in the form of the word: To in-*form* implies that form is introduced into something. A communicative act is realized by in-forming a physical medium with some pattern. A person's understanding of the world as well as that person's social relationships are in-formed in the act of interpreting the message. Human communication must include the cognitive processes by which a message is interpreted and a person's understanding is informed as well as the social processes within which communicative acts have meaning. Any communication must include a physical transmission process, but communication among humans includes much more than physical transmission.

## Signal Transmission

The physical process of creating, preserving, transmitting, and reproducing a signal is the domain of the mathematical theory of communication, also known as *information theory* or *signal*

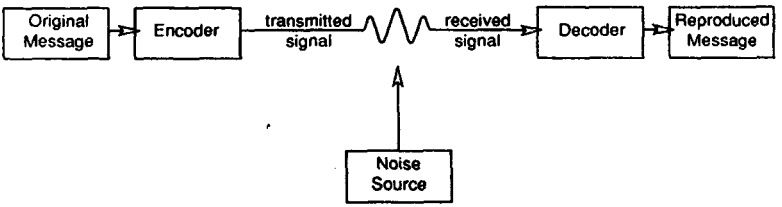


**Figure 2.1.** A Signal Transmission System
SOURCE: Based on Shannon (1949, p. 34)

*transmission theory.* From the perspective of signal transmission theory, the fundamental problem of communication is that of reproducing, at a particular time and place, a message selected at another time and place (Shannon, 1949, p. 31). The process of signal transmission is illustrated in Figure 2.1.

*Encoder and decoder.* The encoder and decoder transform patterns of one type into patterns of another type. For example, the telegraph key encodes each letter of the alphabet into a series of clicks, and the telegraph receiver decodes each series of clicks with a letter of the alphabet. The encoder and decoder come as a matched pair: The decoder ordinarily performs an operation exactly the inverse of the encoder, so as to reconstruct a message as nearly identical to the original as possible. (Shannon, 1949, uses the word *transmitter* for encoder and *receiver* for decoder.) A good example is the telephone, in which the mouthpiece transmits exactly those signals the earpiece is designed to receive.

*Message, code, and element.* In ordinary usage, *message* implies a complete unit of discourse, capable of being transmitted and interpreted independently. A message might be a sentence, a telegram, a letter, or Dickens's *Bleak House.* Each message is composed of smaller units: phrases, words, phonemes or syllables, and letters. The smallest units defined by a code, the units of which messages are composed, may be labeled *elements.* The letter "j" is an element of the English alphabet; dots, dashes, and pauses are elements of Morse code.

The entire set of elements used to form messages in a code can be defined as the *code set.* A *code* can then be defined as a set of

rules for mapping one code set onto another. For example, the system of standardized spelling in English maps phonemes and words onto letters of the English alphabet. A dictionary qualifies as a code inasmuch as it maps letters onto a set of words and words onto a set of definitions.

The original message is composed of elements drawn from the *source code set* and the message is encoded into the *target code set*. For a transmitter using Morse code, the alphabet is the source and the three-element set {dot dash pause} constitutes the target. The transmitter encodes from the source to the target code set and the receiver decodes from the target to the source code set: Decoding is nothing but encoding with the target and source reversed. The person (or thing) responsible for selecting or constructing the original message is the *originator*, and the person or thing that perceives and interprets the reproduced message is the *interpreter*. (Shannon, 1949, uses the single word *source* to indicate both the source code set and the originator of the message.)

A message is a string of elements that can be transmitted independently as a complete specification of some element or a string of elements from the source (see below). The pattern "str" includes three elements but is not a message in English: the pattern "strong" includes six elements of the English alphabet, organized according to the code of American English spelling, to form a message that specifies a particular element of the English vocabulary. According to this definition, the longer string "I like my coffee strong" would also be considered a single message. To consider "I like my coffee strong" but not "millstone fees I my strong" to be recognizable as a "message" requires that the definition of "message" incorporate principles of recognition external to the code, in this case, a natural language by which acceptable word sequences may be distinguished. For the purposes of signal transmission theory, the distinction is important only if those principles of recognition are incorporated into the design of the code (see Chapter 4). Each time an element of some set appears in a message, it constitutes a *datum*.[1] *Element* refers to membership in a set and *datum* refers to part of a message.

*Encoding and decoding.* Encoding and decoding imply a process wherein elements from one code set are matched in uniquely

determined combinations with elements from another code set. A combination of elements from the target code set that specifies an element or combination of elements from the target code set is a *message unit.* For example, Morse code maps letters of the alphabet, numerals, and punctuation onto the three-element set {. - _} dot, dash, and space. In conventional Morse code, "-._" and "- - -_" are message units but "- - - - - - - - - - - -" is not a message unit because it is not defined; it does not map onto or specify any element from any other code set.

Messages may be encoded into a medium with different time-space properties, as when a pattern (e.g., speech or music) is encoded from sound waves into radio waves that can be transmitted over great distances or into magnetic patterns on tape that can be stored for years and replayed dozens of times. A carefully designed code can also alter the statistical properties of the signal so as to improve the rate and/or accuracy of transmission. The primary consideration in designing a code is to match each element in the first code set with some combination of elements in the second code set that will identify the first element precisely to assure accurate recovery during decoding.

Although decoding is the inverse of encoding, a code need not be symmetrical: It is possible for a code to map a given element from the source onto more than one combination of elements from the target code set. For example, English spelling encodes a single element of vocabulary as either *practise* or *practice* and English pronunciation encodes the word spelled *aunt* as "ahnt," "awnt," or "ant."

*Distortion and noise.* The received signal may differ from the transmitted signal in three ways: Elements of the signal may be lost; elements may be transformed into other elements; and elements may be added to the signal. If the difference is systematic and predictable, so that a predetermined operation on the received signal will restore it to a form that matches the transmitted signal, the effect is distortion. Some electric typewriters tend to print "-" instead of "e." Editors and others required to read large amounts of typed prose often develop a habit of mentally substituting "e" whenever they encounter "-."

If the difference is unsystematic, random, and unpredictable, so that it is impossible to recover the form of the transmitted signal,

the effect is *noise.* On old manual typewriters, keys sometimes stick, so that letters randomly fail to print at all. On some models of typewriters (especially when used by amateur typists), any letter can stick: It is impossible to tell what was intended. Once it has been detected and analyzed, distortion can be corrected by adjusting the decoding process; because it is completely random, noise cannot be detected with certainty and can never be totally corrected.

Noise in signal transmission theory can describe silence as readily as sound. The defining characteristic is that, "if the channel is noisy it is not in general possible to reconstruct the original message or the transmitted signal with *certainty* by any operation on the received signal" (Shannon, 1949, p. 66, italics in the original). In this sense, the leaf-blower outside the window or the students gossiping in the back row of the lecture hall are noise only if they randomly obliterate or alter parts of the signal someone is trying to receive and interpret. A flaw in the public address system that randomly cuts out and substitutes silence for parts of the lecture is a source of noise in exactly the same sense as the occasional feedback squawk that overpowers the lecturer's voice. An echo that adds "uh!" at the end of every phrase is distortion, because it is systematic and the original signal is easily recoverable. The world includes a lot of unwanted messages (leaf-blowers, junk mail, TV commercials, snippets of somebody else's conversation in a long-distance telephone connection) that are not noise in the sense of signal transmission theory, because they do not alter a transmitted signal in a way in which it cannot be recovered.

*Efficiency and accuracy.* A primary objective of signal transmission theory is to understand the constraints on efficient and accurate signal transmission: Given that any physical transmission system can carry only a certain amount of messages, what determines the maximum amount that can be carried with reasonable accuracy? Shannon's contributions to communication theory included developing rigorous ways to estimate the maximum capacity of a system with known physical characteristics and to calculate the trade-off between accuracy and rate of transmission when the system is plagued by noise. Shannon showed that the solutions to these questions of efficiency and accuracy are

implicated in the power of a code to specify elements of any source code set: How many elements of the target code set are required, on average, to specify each element of a typical message from the source code set?

*Variety.* It would be impossible to form messages using a one-element code set: Even a code based on tapping on a tabletop requires intervals of silence between taps. Because a minimum code set must have at least two elements, a basic requirement for communication is variety. Variety increases as a function of the number of elements in a set and as a function of the frequency with which each element is used. For example, if a code used the 26 letters of the alphabet in such a way that 99% of the observations were either "j" or "k," the code would have little more variety than a code with only two elements. Shannon showed that the power of a code is a function of the amount of variety among the elements of its code set as they appear in a typical message. This idea of variety, along with the formula for computing the amount of variety in the way a code uses its code set (see Chapters 3 and 4), has many potential applications beyond estimating the power of a code and the efficiency of a signal transmission system.

## Interpretation

It is a widespread practice to speak and write about the processes of interpreting communicative events as *decoding*, a usage that is probably justified for many routine interactions. If someone nods in reply to a question, the nod is decoded as "yes." "Yes," whether encoded as a spoken phoneme or as a head nod, is decoded as assent or affirmation. "I'm going home now" can ordinarily be interpreted by simply decoding each word into a commonly implied person, action, destination, and time while decoding each word position into a grammatical function.

However, much of human communication is indirect and implied or metaphoric (Grice, 1975; Lakoff & Johnson, 1980; Richards, 1936; Sperber & Wilson, 1986, 1987; Terwilliger, 1968). Phrases commonly regarded as having a fixed, codelike meaning (e.g., "his temperature is rising") are frequently based on systems of metaphor and cannot always be decoded in a straight-

forward way. Even when communication is confined to a precise code, humans can and do distort, expand, and subvert the code to express novel meanings. As simple a phrase as "I'm going home now" may be quite ambiguous, because "home" may refer to a nation, state, city, house, social group, or even a frame of mind. "Now" literally encodes "the action is in process as I speak" but often implies "the action will commence as soon as we've finished saying good-bye and I've buttoned my raincoat."

*Relevance and context.* Sperber and Wilson (1986, 1987) suggest that actions are classified as communicative (i.e., as messages) either because they fall into a conventional category such as speech vocalizations or gestures or because they are performed in a way that commands attention and can only be explained as communication. Any such action is a claim on another person's attention, an implicit promise that it is worth the perceiver's effort to figure it out. Whenever any action (linguistic or otherwise) cannot be readily explained, the perceiver can be expected to interpret the action by seeking its relevance in the context of what has already happened or might be expected to happen.

Context includes whatever the originator of a message might expect an interpreter to know. Context would thus include ideas already introduced in the conversation, readily perceived aspects of the physical environment, or ideas that may be expected to be recalled from memory. A claim of relevance is a claim that the message will combine with and alter the context in some way, for example, by adding new facts or ideas or by changing previously available ideas. Interpretation requires discovering the relevance of a message and may include expanding or changing the interpreter's understanding of what the context is. For example, the interpreter may have thought the context of a conversation with an old acquaintance was "the good old days," until a remark about "making sure your loved ones are provided for" invokes the very different context of "selling life insurance." Once the new context has been invoked, the relevance of previous parts of the conversation may change in a radical way. For example, questions about one's career development may be reinterpreted as an attempt to gauge the likely size of an expected sales commission rather than as evidence of personal interest.

*Changing the context.* Messages change contexts in several ways. The interpreters' awareness may be changed slightly, by introducing new ideas in a context already salient, or their awareness may be changed profoundly, by calling to mind entirely new contexts and suggesting previously unsuspected linkages among these contexts. Moreover, the social context of relationships linking message originator and interpreters may also be changed, as new ideas are made common and as mutual expectations for future interactions are altered. Each of these effects is included in what we mean by *information.*

## Data, Relevance, and Information

It appears that animals, like computers and other machines, communicate by means of codes that link stimuli to responses in a readily predictable way (Wilson, 1971). The messages are made up of elements from a set of stimuli, paired with particular responses by some code (see Figure 2.2). Each stimulus counts as a datum only inasmuch as it is paired by the code with some particular response.

As with insects and machines, human communication sometimes relies on stimulus-response codes; for example, the infant's smile elicits the parent's nurturing response. Thus, even in human communication, the concept of information includes coded data. However, even the most animal-like codes can be used to communicate something more or even entirely different than the encoded idea or response. Young children learn to create the smile for its effect—and adults learn to discount such uses of the code, to ask somewhat cynically, "What's the little devil up to now?" The smile has become an unexplained behavior that requires interpretation in some contexts, for example, the context of a newly filled cookie jar. The older child even learns to use the smile ironically, to express the *opposite* of its coded idea. In response to an overweight parent's trip to the cookie jar, the smile becomes a smirk.

Humans smile, frown, growl, emit cries of pain or rage, make the erotic gestures of courtship and the hostile gestures of aggression. Sometimes humans use these codes to express deliberately what the gestures would ordinarily imply; sometimes (as when a smile becomes a smirk) the gestures are used metaphorically or ironically to express ideas of an entirely different sort.

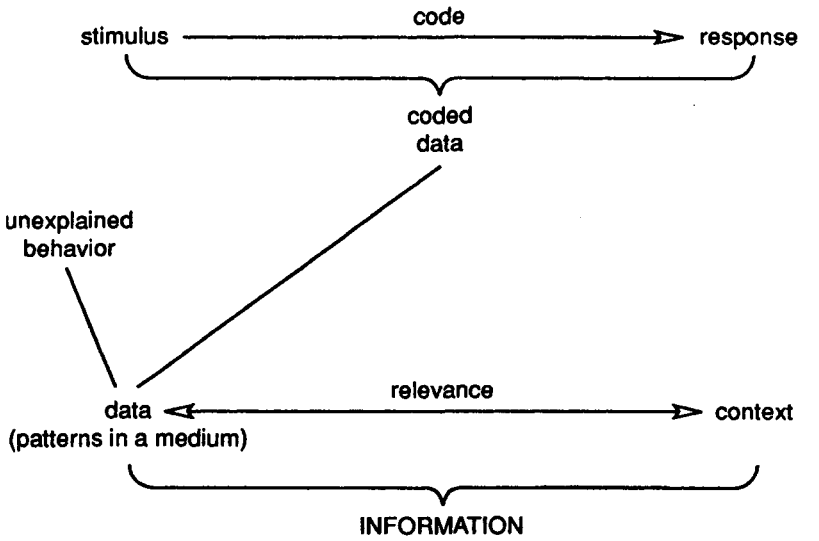


**Figure 2.2.** Data and Information

Humans also communicate by performing otherwise meaningless actions in a noticeable way; for example, exaggerated gestures or prolonged silences imply communicative intent. Although they are part of no code set, these gestures become data for human beings. The child may hold up an empty plate, or stare at the cookie jar, or stare at another child who is eating a cookie. The child is using his or her body as a medium; the gestures are patterns that, once perceived as data, become relevant in the context of what the adult knows about the relationship between children and cookie jars. As relevant data, these gestures become information; they tell or inform the parent about the child's desire for a cookie. Combined with the parent's response, they also sustain or restructure the parent-child relationship.

For humans, data include any patterns in a medium that can be interpreted as relevant to some context. Data gain meaning, they become information in the human sense, through interpretation. Information includes both data and the relevance of data in some context. We can rarely be quite certain about the contexts that will be intended by a communicator or understood by

an interpreter; in human communication, unlike animal or machine communication, the idea of information inevitably includes a certain degree of ambiguity.

Pattern and relevance are not alternative forms of information; they are but two ways of describing the same information. Pattern has to do with mediation and describes the relationship between physical form and how much is (or can be) indicated by a message. Relevance has to do with meaning and describes the relationship of patterns and whatever patterns indicate to the cognitive environments of the originator and perceiver of a message. Patterns with communicative potential are data, and data with relevance are information. Data may have more or less relevance, but events must be at the very least relevant to some communicative interaction to be accepted as data. The concept of information includes the concept of data and, conversely, the concept of data can be defined only within the concept of information.

Information requires pattern, and pattern requires both variety and structure. The concepts introduced in the first part of Chapter 2, having to do with encoding and transmission, provide the basis for a discussion of variety in Chapter 3 and structure in Chapter 4. The concepts introduced in the second part of Chapter 2, having to do with interpretation, provide the basis for a discussion of the relationship between message and the context of communication, beginning in Chapter 5.

## Note

1. *Data* is the plural of the Latin word *datum*, just as *media* is the plural of *medium*. *Datum* comes from *dare*, "to give": literally, a datum is something that is given.

# 3. *Measuring the Variety of a Set*

Included in the broader question of how much can be communicated in a situation is the problem of the capacity of the systems

available for transmitting messages. *Transmission* here refers generally to the problem of reproducing messages at another time (i.e., representation and storage) as well as the problem of reproducing messages at another place (i.e., transmission as it is commonly understood).

The problem of transmission capacity can be restated as a problem of calculating the power of a code to specify elements of one set by matching them with elements of another set. Transmission is accomplished by encoding the message into a target code with the desirable characteristics of duration and/or motion: Once the power of the target code is known, the capacity of the transmission system can be computed. Computing the power of a transmission system requires measuring the variety of elements in the source set—from which come the messages to be transmitted—and the variety of elements in the target set—into which the messages are encoded for transmission.

A clear understanding of ideas such as variety, transmission capacity, efficiency, and redundancy requires familiarity with a few mathematical expressions. It is particularly important to understand the logic underlying the formula for H. This chapter explains the most important mathematical expressions and their relationship to encoding and signal transmission. The reader may find it useful to skim lightly over much of the material in Chapter 3 the first time through, then refer back to the more difficult sections for closer study as needed.

## The Power of a Code

The conventional system of counting and numbering things uses 10 digits; the English alphabet uses 26 letters; a typical adult has a vocabulary of many thousands of words. In English spelling, many thousands of words can be represented by combining a few of the 26 letters plus a space or punctuation to mark the end of a word. In Morse code, each letter can be represented by a combination of one to four dots and dashes plus a space.

Because fewer symbols are required to communicate a word when the symbols are letters of the alphabet than when they are dots and dashes, a code that uses the alphabet is likely to be more powerful than a code based on dots and dashes. However,

the power of a code is not a simple function of the number of elements in the code set N. It is also a function of how efficiently these elements are used in the code. The code of English spelling uses only a fraction of the possible combinations. For example, of 51 possible two-letter combinations that include "t," 41 are ruled out by the requirement that every word have at least one vowel.

The power of a code is a function of the variety of elements and the variety of patterns formed by combining these elements. The variety of English spelling, and consequently the number of things that can be communicated by combining a certain number of letters in English, is in turn a function both of the number of letters and of the efficiency with which these letters are combined.

## A Basic Metric: The Binary Set

Comparing the communicative power of various codes requires a common unit of comparison. It is convenient to use a two-element (binary) code set as the basic unit of comparison, for at least two reasons. First, two is the smallest number of elements that can possibly be used to form a pattern, hence to communicate anything. Second, it happens that a number of gadgets of interest to communication scientists, including the digital computer, are designed around binary sets. Accordingly, in this chapter, the logic of a binary code set will be developed as a basic metric for representing and comparing codes that use larger sets. The result is a formula that can be used to express the amount of variety in the elements of any set—such as messages produced by a code, responses to a multiple-choice questionnaire, students' choices of college majors, parents' choices of first names for their children, or customers' selections from a restaurant menu—and to compare the amount of variety in any two or more sets.

*Encoding into the binary set from a four-element set.* Human beings have long been fascinated with questions of efficient coding; many popular puzzles and parlor games (e.g., "twenty questions") are constructed around the problem of efficiently encoding information. Imagine such a puzzle, a guessing game in which you are given the task of communicating the content of a series of cards to a friend by stringing beads on a wire. The

| Code Set 2 | | Code Set 1 |
|---|---|---|
| a | = | rw |
| b | = | rr |
| c | = | ww |
| d | = | wr |

**Figure 3.1.** Translating from Four Elements to Two Elements

beads are identical except that half are red and half are white; that is, they are all drawn from the two-element set {r w}. The cards will be drawn at random from a large deck of alphabet cards, each imprinted with a letter drawn from the set {a b c d}. After you have been shown a series of cards, your partner will guess which cards you were shown. You will be awarded points for each correct guess, but you will lose points for each incorrect guess. Because you will also be penalized for the number of beads you use, the game puts a premium on using as few beads as possible. (Many real-life communication problems are subject to similar efficiency constraints. For example, classified ads are priced by the word and long-distance telephone calls by the minute.)

Before the game starts, you and your partner must construct a code that combines the beads in various ways so as to communicate what you have seen, using as few beads as possible. The task can be generalized as the problem of constructing a maximally efficient code for specifying messages from a four-element set using elements from a binary set. One possible solution, shown in Figure 3.1, requires only two beads for each card.[1]

If you are shown the card series "abddca," you would arrange the beads so that your partner would see "rwrrwrwrwwrw." To signal your partner, you will require twice as many beads as the number of cards you are shown. Within these constraints, it is impossible to arrive at a more efficient coding system.

How much does a single card communicate? The beads are an example of a binary code: Each bead communicates one bit, the smallest unit of communication possible. It requires two beads to specify a single card, and a single card can specify two beads. If a code is developed to express messages from some other set (for example, to label greeting cards in an inventory), a code using {red white} beads would require twice as many elements as a code using cards imprinted with {a b c d}. Because a single card

| Code Set 3 | Code Set 1 |
|---|---|
| 1 | rrr |
| 2 | rrw |
| 3 | rwr |
| 4 | rww |
| 5 | wrr |
| 6 | wrw |
| 7 | wwr |
| 8 | www |

**Figure 3.2.** Translating from an Eight-Element to a Two-Element Code Set

specifies as much as two beads, we can say that a code using the four-element set is twice as informative as a code using the two-element set, and the typical element from the four-element code conveys two bits.

*Encoding into the binary set from larger sets.* Now imagine a higher-level game, using a deck of cards in which each card is imprinted with an element chosen from the set {1 2 3 4 5 6 7 8}. How few of the binary code beads could you use to communicate about the eight-card deck? Figure 3.2 shows one possible solution.

Although there are twice as many elements in Code Set 3 as there are in Code Set 2, only one more element from the binary set is required to specify an element from Code Set 3. Code Set 2 requires two elements, and Code Set 3 requires three elements. Conversely, any string of N beads from Code Set 1 can be represented by a string of N/3 cards from Code Set 3. A typical element from Code Set 3 conveys as much as three typical elements from the binary set (three bits).

This result can be readily generalized to sets with larger numbers of elements. For example, it is a simple task to construct a code in which only four beads can be arranged in sequences that specify any card from a deck of 16 distinct cards (a 16-element set). All that is required is to use a fourth bead to specify whether the observed card is among the first or the last eight; the code shown in Figure 3.2 may then be used to identify which of the eight is intended. For a 32-element set, a fifth bead is added, to specify whether the card is among the first or last 16, and so on.

*$H_{max}$: The maximum amount of variety.* The number of binary elements required to specify an element from a target set increases by 1 each time the size of the target set doubles. Conversely, each time the number of beads (binary elements) in a string is increased by one, the power of the code—its ability to specify an element from some other set—is doubled. The power of a code using strings of H elements from the binary code set can be expressed in terms of N, the number of elements the code can uniquely specify, by multiplying the number of elements in the target code set (two) times itself H times:

$$2^H = N \qquad [1]$$

Conversely, when N—the number of elements in the code set—is already known, the number of bits represented by or required to represent the typical element of the set can be computed by taking the logarithm of N, obtained by solving for H:

$$H = \log_2 N \qquad [2]$$

When base 2 logarithms are used, as in Equation 2, H expresses the average number of binary digits (bits) required to specify any one element of a code set with N elements, *if* those elements appear with equal frequency. As will be seen, if a code does not use its elements with equal frequency, its ability to specify elements of any other code is reduced: The maximum value of H is achieved when all elements appear with equal frequency. Accordingly, the quantity expressed in Equation 2 is often labeled $H_{max}$ to denote the maximum value of H that can be achieved by a code using N elements.

The logarithms are calculated to the base 2 to express the power of a code in terms of binary digits. There is no inherent reason why H could not be expressed in terms of a "basic" code set of any other size, M, in which case the value of H would be calculated in terms of base M logarithms.

$$H = \log_M N \qquad [3]$$

In Equation 3, M is the number of elements in the target set and H is the number of elements from the target set required, on

average, to specify any element from a source set with N elements. In the binary code, M = 2.

Many of the code sets of interest to communication researchers contain elements that are not evenly distributed. (The letters of the alphabet, for example, appear in written English with a very uneven distribution.) Code sets with an unequal distribution of elements will be discussed in the next section, but first it is necessary to convert Equation 2 from an expression based on the number of elements in a set, N, to an expression based on the probability of observing any particular element.

First, using

$$\log 1/N = -\log N \qquad [4]$$

the expression in Equation 2 is equivalent to

$$H = -\log_2(1/N) \qquad [5]$$

As long as all of the elements are evenly distributed, the probability p(i) of observing the $i^{th}$ element is equal to 1/N for any element i, where $1 \leq i \leq N$, and so

$$H = -\log_2 p(i) \qquad [6]$$

The expression given in Equation 2 describes the average number of bits that can be specified by any element from a set in which the elements appear with equal frequency. The expression in Equation 6 describes the number of bits that can be specified by a particular element i that appears with a probability p(i). As long as the elements of the set are distributed with equal probability, p(i) = p(j) = 1/N for any two elements i and j, and the expression in Equation 2 is equivalent to the expression in Equation 6. Thus, when the elements of a code set are used with equal frequency, each element in the set has a power to specify elements of the binary set that is equal to the power of every other element, and the value of H for each element in the set is identical to the average value of H for the typical element.

Note that the minus sign is strictly a mathematical consequence of the transformation from N to 1/N; it has no other significance. Some writers have suggested that the minus sign is an

arbitrary convention, introduced to give H a positive value. Although the expression for $H_{max}$ in Equation 5, like the expression in Equation 2, certainly *does* have a positive value, there is nothing arbitrary about it. The formula $H_{max} = \log_2 N$ follows from the definition of H as the power of a code to specify elements of the binary code set, and the minus sign follows from transforming the formula from an expression of the total number of elements in the code set (Equation 2) to an expression of the probability of observing any particular element (Equation 6).

## Codes with Unequal Distribution of Elements

Thus far, the discussion has proceeded according to the assumption that a code uses the elements in its code set with equal frequency. For example, the discussion of Code Set 3, the set {1 2 3 4 5 6 7 8}, assumed that the deck included an equal number of cards with each symbol; each element appears with a probability p(i) = 1/8. However, many of the most interesting codes for the study of human communication do not use the elements of their code sets in such an equal way. The code of standard English spelling, for example, uses the elements of the English alphabet in such a way that they are very unequally distributed: "e" and "t" appear quite frequently but "x" and "z" are rare.

*Encoding unequally distributed elements.* Consider a simpler example, using the same eight elements as in Code Set 3. Suppose messages are constructed using the eight elements in Code Set 3' according to the following rules: (a) Every string is exactly two elements long; (b) the first element is always either 1 or 2; (c) if the first element is 1, the next element is either 3 or 4; (d) if the first element is 2, the next element is drawn from the subset {5 6 7 8}. These rules will produce a distribution of elements as follows: 1 and 2 are both twice as likely to show up as either 3 or 4 and four times as likely to show up as 5, 6, 7, or 8; both 3 and 4 will then be twice as likely to show up as 5, 6, 7, or 8. Note that an infinite number of sets of rules can produce the same probability distribution of elements. For example, consider a code that might be constructed by altering rule a to require that every string have an even number of elements and by changing "the first element" in

| Code Set 3′ | probability | Code Set 1 | Bits |
|---|---|---|---|
| 1 | 1/4 | rr | 2 |
| 2 | 1/4 | rw | 2 |
| 3 | 1/8 | wrr | 3 |
| 4 | 1/8 | wrw | 3 |
| 5 | 1/16 | wwrr | 4 |
| 6 | 1/16 | wwrw | 4 |
| 7 | 1/16 | wwwr | 4 |
| 8 | 1/16 | wwww | 4 |

**Figure 3.3.** Translating from an Eight-Element Code Set with Unequal Distribution to a Two-Element Code Set

rule b to "every odd-numbered element." The distinction between rules of a code and the resulting statistical distribution of elements in a typical message will be taken up in Chapter 4.

How many beads from the binary {red white} set will be required, on the average, to specify a card from Code 3′? About half the time, the card to be specified will be either a 1 or a 2, so the first element in the string tells whether the card is lower than 3. If the answer is yes, the second element will identify it as either a 1 or a 2. In this way, a string of only two beads will suffice to specify either of the two most common elements.

If the card is not lower than 3, then it is most likely to be either a 3 or a 4, so the second element in the string indicates whether the card is lower than 5. If it is, one additional element will suffice to specify whether the card is a 3 or a 4. For the second most common set of elements, a string of three beads is required. If the card turns out not to be lower than 5, then it is equally likely to be any of the four least common cards. Two beads are needed to determine that the card is one of the four least common elements; an additional two beads will specify which it is (Figure 3.3).

Because the number of elements in a string is no longer constant, the code has to be constructed in such a way that it is evident where one string ends and the next begins. In Figure 3.3, each two-bead string begins with a red bead, and each three-bead string begins with a white bead followed by a red bead.

$H_{obs}$: *The observed amount of variety.* The number of bits represented by each element of the code illustrated in Figure 3.3 can

be calculated by substituting the probability associated with each element for p(i) in Equation 6:

$$\text{Elements 1-2: } H_2 = -\log_2(1/4) = 2 \quad [7]$$

$$\text{Elements 3-4: } H_2 = -\log_2(1/8) = 3 \quad [8]$$

$$\text{Elements 5-8: } H_2 = -\log_2(1/16) = 4 \quad [9]$$

The *average* number of bits per element can be computed by multiplying the number of bits specified by each of the eight elements by the frequency with which that element occurs in a "typical" message and summing the products, thus weighting the value of each element in the set by the frequency with which it is observed:

$$H = -\Sigma p(i) \log_2 p(i) \quad [10]$$

where the summation is over i = 1 through N. The formula in Equation 10, applied to the distribution shown in Figure 3.3, yields the following results:

$$\text{Element 1: } H_2 = -1/4\log_2(1/4) = 1/4 \times 2 = 1/2 \quad [11]$$

$$\text{Element 2: } H_2 = -1/4\log_2(1/4) = 1/4 \times 2 = 1/2 \quad [12]$$

$$\text{Element 3: } H_2 = -1/8\log_2(1/8) = 1/8 \times 3 = 3/8 \quad [13]$$

$$\text{Element 4: } H_2 = -1/8\log_2(1/8) = 1/8 \times 3 = 3/8 \quad [14]$$

$$\text{Element 5: } H_2 = -1/16\log_2(1/16) = 1/16 \times 4 = 1/4 \quad [15]$$

$$\text{Element 6: } H_2 = -1/16\log_2(1/16) = 1/16 \times 4 = 1/4 \quad [16]$$

$$\text{Element 7: } H_2 = -1/16\log_2(1/16) = 1/16 \times 4 = 1/4 \quad [17]$$

$$\text{Element 8: } H_2 = -1/16\log_2(1/16) = 1/16 \times 4 = 1/4 \quad [18]$$

The sum of Equations 11 through 18 is

$$H = -\Sigma p(i)\log_2 p(i) = 2.75 \qquad [19]$$

Each of the most common elements, 1 and 2, represents only half as many bits as one of the least common elements, 5 through 8, but both 1 and 2 appear four times as often and thus contribute twice as much to the average value, H. Because Code Set 3′ has less variety than Code Set 3, each element has (on average) less power to specify elements of any other set, and a message in Code 3′ will be slightly longer than an equivalent message in Code 3. The rules by which Code 3′ is constructed underuse half the elements in its code set and reduce the power of the code accordingly.

Because Equation 10 is based on observed frequencies, the value of H calculated by Equation 10 is often labeled $H_{obs}$. When the N elements are equally distributed, so that p(i) assumes a constant value equal to 1/N, note that Equation 10 is equivalent to Equation 6 and $H_{obs} = H_{max}$, because

$$\begin{aligned} H_{obs} &= -\Sigma[(1/N)\log_2(1/N)] \qquad [20] \\ &= -(N/N)\log_2(1/N) \\ &= -\log_2(1/N) \\ &= H_{max} \end{aligned}$$

However, when the N elements are not equally distributed, $H_{max} > H_{obs}$. The difference, $H_{max} - H_{obs}$, suggests a way of describing the efficiency of a code, a result that has proven useful in electronic communication engineering, and also suggests several interesting ways of analyzing natural communication events. For Code 3′, $H_{max} - H_{obs} = 3.00 - 2.75 = .25$ bit (see Chapter 4). Note that, when elements are unequally distributed, the value of H for the typical element is an average and means something different than the value of H for any particular element.

## What H Measures

H measures the amount of variety in any set of observations, including messages in some code. H is usually expressed as the

number of bits of a binary code required, on average, to specify one element of the described set or, conversely, as the number of elements of a binary set specified by an average element of the described set. H expresses the capacity of a code to represent or transmit messages, regardless of what the messages are, what code they are in, or how they came into existence. In effect, H measures the capacity of a code to represent or transmit *data*. H is often called a measure of the *information* capacity of a code (or of a transmission circuit using a code with certain statistical characteristics). Because meaning—any relationship of data to a context or purpose—is excluded from signal transmission theory, within the context of signal transmission theory, saying that H measures data capacity is equivalent to saying that H measures information capacity. The distinction between the two statements only becomes important when relevance is considered, as it is in human communication.

*Measuring H for a particular message.* The logic of H can be extended to develop a measure of the capacity of a particular message, that is, a string of elements in some code. The contribution of each element to the power of a code as measured by Equation 6 can be treated as the data capacity of that particular element. If the data capacity of each element is counted every time it appears as a datum in a particular message, the sum can be used to estimate the data capacity of the total message.

Some care is advisable in how the results are interpreted. First, the value of H for each datum must be calculated on the basis of the distribution of elements in the code, not on the basis of the distribution of data in any particular message. It is especially important in dealing with a particular message to remember that probabilities always refer to the general case, prior to any observation, and cannot be calculated for a particular message that has already been observed. Second, the question of external structure, the relationship of the message to the overall context, should be segregated from the analysis. This will be facilitated if the value of H is labeled either "H" or "the data capacity of the message." Especially if the message is in a natural human language, the word *information* should be avoided because it too readily leads to confounding the statistical characteristics of a code with the cognitive and social processes of communication.

## Some Useful Characteristics of H

H has several useful mathematical characteristics. Because H represents variance on a nominal scale, it is applicable to any data without parametric assumptions. Because it is expressed in logarithms, H creates a ratio scale: It has a true zero point, and all mathematical operations can be performed on it. For example, 2 bits is twice as much as 1 bit in exactly the same sense that 30 bits is twice 15 bits. H can easily be converted by multiplication and division from one code base to another (for example, from a binary to a hexadecimal base). Because values of H are subject to normal arithmetic operations, the power of any code can be compared either with the "ideal" maximum power of a code using its own code set or with a code based on some other code set. The efficiency with which a code uses its code set can be readily computed and compared with the efficiency of other codes (for a detailed description of statistical techniques using H, see Krippendorff, 1984).

## Note

1. At first glance, it might appear that a more efficient code could be constructed by using two strings of one bead each, "r" and "w." However, such a solution makes no provision for determining where one message ends and another begins. In general, codes that employ strings of varying length must either employ a device, such as the space used after each word in English spelling, or restrict the allowable short sequences in such a way as to use position to mark the end of a string.

# *4. Redundancy: How Structure Affects Variety*

H measures the variety among the elements in any set. Chapter 3 introduced two ways to calculate H, $H_{max}$ and $H_{obs}$. In Chapter 4, these two quantities will provide a basis for estimating the

degree to which structural constraints on the way a code uses its elements reduces the amount that can be communicated by using the code. $H_{max}$ measures the maximum variety that would be observed if all elements appeared with equal probability, and $H_{obs}$ measures the variety of actual observations: The difference between these two quantities is *redundancy*, the reduction in the efficiency of the code due to constraints on the use of its elements. Redundancy is useful for analyzing the trade-off between efficiency and the structure generated by the rules that constitute a code. The structure of a code can constrain its use of elements in a number of ways: Before H can be used to estimate the communicative capacity or potential of a code, or of a particular message, it is necessary to understand more fully how the structure of a code constrains the variety of its messages.

## Code Efficiency

In Chapter 3, the amount of variety produced by using a code, H, was calculated for two different codes, both drawing on the eight-element set {1 2 3 4 5 6 7 8}. One of these codes has an average of H = 3 bits per element; the other has an average of H = 2.75 bits per element. The difference is due to the fact that the elements in the code set are equally likely to appear in a message formed by the first code, but they are not equally likely to appear in a message formed by the second code.

Given any code that includes a fixed number of elements, N, the maximum value of H will be obtained when all elements are equally probable. That does not mean that every message will contain an exactly identical distribution of the elements in the code, because H is always calculated for the distribution of elements in a *typical* message. It only means that, in a sufficiently large random sample of messages drawn from the code set, the distribution will tend toward equal probability.

## Redundancy

The degree to which a code deviates from its theoretically maximum efficiency is defined as the *redundancy* of the code. In

everyday language, *redundancy* also refers to the *repetition* of a particular message, as, for example, when we explain something more than once or when a friend tells about last summer's trip to Europe for the fifth time. In signal transmission theory, *redundancy* refers only to inefficiencies in the way a code uses its elements and always refers to general characteristics of the code, never to characteristics of a particular message. Message repetition can serve some of the same functions as code redundancy, for example, to minimize errors, and repetition can be built into the structure of a code. But repetition can also be used to communicate new ideas, as when an exasperated parent repeats a request to emphasize that "this is a demand, not merely a request" or when a hopeful lover repeats a declaration of love to emphasize the sincerity of the declaration.

For example, redundancy is a major tool used by cryptographers to break codes. (A simple application of cryptography is in solving the cryptograms that appear in the entertainment sections of many Sunday newspapers.) English uses 26 letters plus a space and a few punctuation marks, but these elements are used with quite unequal frequency. In solving a cryptogram, the first step is to try to identify the most common events. In English, these are likely to be associated with "e," "t," and the space between words. A second useful step is based on the fact that the most common word in English is *the*. The cryptographer uses this fact by looking for repeated three-character sequences in which the first and third characters are among the most common. Including the space allows the cryptographer to expand the search to five-character sequences including a repeated beginning and ending: " t_e " is easily identifiable even in short cryptograms.

The facts that "t" and "e" are observed more frequently than "x" and "q" in everyday English are examples of *distributional redundancy*. The facts that "t" is often followed by "h" then "e"; "q" is usually followed by "u"; and every word includes at least one vowel are examples of *structural redundancy*. *Structure* refers to the manner in which the rules of the code organize its elements into messages, and *structural redundancy* refers to the reduction in communicative power that is associated with structure. In English, these rules govern spelling, the importation of words from other languages, the co-occurrence of sounds

within a word, and sentence construction. Rules of spelling and importation of words are responsible for the distributional frequency of the letter "e"; rules of sentence construction are responsible for the distributional frequency of the word *the*.

## Structure

The structure of a code can be described in terms of its *functional structure*—the rules or processes by which it assembles messages—or in terms of its *statistical structure*—the conditional relationships among elements that are observed as a result of the operation of these rules. The statistical structure of a code is expressed as the *conditional probabilities* associated with each element. *Conditional probability* refers to the probability of observing any particular element, based on what has already been observed. For example, the conditional probability that the next letter in a word will be "e" is higher if the previous three letters were consonants than if they were vowels.

*Internal structure.* Relationships among the elements of a code, as produced or influenced by the rules for assembling elements into messages, constitute the *internal structure* of the code. The internal structure of a code can be defined in two ways: Functionally, structure is the set of rules by which the code combines elements of its code set to form messages. The rules may be formally stated and applied, as in an artificial language or code, or the rules may be abstracted and inferred by observing usage, as in a natural language.

Statistically, structure can be identified by the conditional relationships among elements of a typical message. For example, English spelling specifies that "q" is followed by "u"; these rules produce a conditional relationship in which the probability of observing "u" is very near 100% whenever a "q" has just been observed. The conditional relationship in which "e" is likely to be observed within one element after "t" results from several inferred rules, including the requirement that each word include at least one vowel, and the use of a "silent e" after a consonant to alter the pronunciation of a preceding vowel.

*External structure.* To communicate anything, the messages formed by the rules of a code must be related in some way to the context of ongoing purpose, perception, and action. No matter how highly structured it is, a code with no relationship to anything else would be literally nonsense. The relationship of messages to context constitutes the external structure of a code. Although these two aspects of structure often interact, external structure is excluded from signal transmission theory. This chapter is concerned exclusively with internal structure; external structure will be discussed in Chapter 5.

## How Internal Structure Contributes to Communication

Because internal structure restricts the code to certain combinations, it unavoidably reduces the power of a code to specify elements of another set. However, internal structure serves several functions that more than compensate for this reduction in power.

*Detecting and correcting errors.* One function of structure is to combat random errors caused by noise in signal transmission. One of Shannon's major contributions to signal transmission theory was to demonstrate that, as long as the amount of noise in a transmission system is less than the capacity of the system, the level of inaccuracy can be brought arbitrarily close to zero by increasing the structure of the code and consequently devoting some of the transmission to error detection and correction.

The simplest form of internal structure is repetition: If the second version of a message differs from the first, the receiver can assume that the difference is due to transmission errors. However, if the repetition is not incorporated into the structure of the code, the second version may be interpreted as a new message. A phrase such as "I repeat," often used in military commands ("Commence firing; I repeat, commence firing"), may be used to mark a repetition as a structural part of the first message rather than an additional message.

A more complex form of internal structure is created by rules restricting the way elements may be combined in a message. Strings that violate these constraints can then be regarded as suspect. One example of such constraints is spelling rules: If we

read "the dpg left a mess on the neighbor's lawn," we know that one of the letters in "dpg" is an error. Because no three-letter word begins with "dp" or ends with "pg," we infer that the "p" is an error. The word is more likely to be *dig*, *dug*, or *dog*. Another kind of structural constraint is grammatical rules. Grammar also contains structure, but either "dig" or "dug" may be used as nouns, and cannot be ruled out on the basis of grammar. "The dug left a mess on the neighbor's lawn" is meaningless, and can be ruled out on the basis of *external* structure. "The dig left a mess" can only be evaluated by looking at the context: is the property near an archaeological site? (see Chapter 5).

*Limits to error correction.* In any transmission, it is possible that more than one element was altered: The transmitted message may have been "the pig left a mess on the neighbor's lawn" or even "the kid left a mess on the neighbor's lawn." Even if the received message looks right—"the dog left a mess on the neighbor's lawn"—we can't be sure that the transmitted message wasn't "the kid left moss in the neighbor's pond." In a world in which every part of the message is subject to random events, no amount of redundancy will assure perfect transmission.

*Detecting communicative signals.* A second function of structure is to identify an event or set of events as a signal to be decoded and interpreted. The concept of a maximally efficient code, in which all elements are equally likely to appear in any position, is central to signal transmission theory. (Except for the fact that zero never appears in the first position, the system of Arabic numerals is an example of a maximally efficient code.) A maximally efficient code has no internal constraints on the use of its elements. No matter what elements have already appeared in a message, each element has an equal probability of appearing next. Each message in a maximally efficient code is indistinguishable from a perfectly random string of events (Pagels, 1988, p. 57).

Consider the problem of detecting possible transmissions from other inhabited planets. Unless it contains a very high degree of internal structure, such a signal could not possibly be detected—neither could it be interpreted if it were detected. If some alien civilization had managed to achieve perfect transmission and reception, and if they capitalized on this

achievement by developing a maximally efficient code, we would have no way of distinguishing their signals from the background chaos.

*Fitting the code to the system.* A third set of constraints on the structure of a code derives from the characteristics of the signal transmission system: Depending on how signals are created and detected, certain combinations of elements may be awkward to produce, transmit, detect, or interpret. For example, the human vocal system has difficulty producing certain combinations of sounds (e.g., the "tongue twister" so popular among young children). The mental processes by which sounds are interpreted also constrain the way elements of a code can be assembled into messages: Rules for spelling and importing words from other languages reflect these restrictions (Stubbs, 1980).

## How Internal Structure Affects H

Structural redundancy is theoretically prior to distributional redundancy, but distributional redundancy is observed before structural redundancy. For this reason, and because the mathematics of distributional redundancy are more readily understood, the discussion will begin with distributional redundancy.

*Distributional redundancy.* The unequal distribution of code elements reduces the power of a code by an amount that may be expressed in terms of $D_{dist}$—the decrease in H as a result of distributional inequalities:

$$D_{dist} = H_{max} - H_{obs} \qquad [21]$$

The distributional redundancy of a code, $R_{dist}$, can then be defined as the ratio of the reduction of H to the maximum value of H for the code set:

$$R_{dist} = (H_{max} - H_{obs}) / H_{max} \qquad [22]$$

$$= 1 - (H_{obs} / H_{max})$$

In Code 3, with an equal distribution of elements, each element represents an average of 3 bits. In Code 3′, as described in Chapter 3, some elements appear much more frequently than others, with the result that each element represents an average of only 2.75 bits.

$$D_{dist} = H_{max} - H_{obs} \quad [23]$$

$$= 3.00 - 2.75$$

$$= .25 \text{ bits}$$

The distributional redundancy is

$$R_{dist} = 1 - (2.75 / 3.00) \quad [24]$$

$$= 1 - .917$$

$$= .083$$

The unequal distribution of elements in Code 3′ decreases its power by .25 bits, or about 8.3% of the power achievable by a maximally efficient code using the same elements with an equal distribution. The distributional inequalities of letters in typical passages of English prose decrease H from about $H_{max} = 4.75$ to $H_{obs} = 4.10$, reducing the power of the English alphabet by about .65 bits, 13.7% of what could be achieved if all letters were used equally.

*Structural redundancy.* Structural redundancy can be expressed statistically in terms of "conditional probabilities," the way in which the probability of observing a particular event is affected by previously observed events. For convenience, the conditional probability of observing event "i" when one or more previous events are known will be labeled $p_c(i)$. The probability of observing event "i" is the sum of the conditional probabilities of observing "i" for each possible combination of prior events, multiplied by the probability of each combination. The probability that "e" will be the third letter of a word can be computed by the product

of the conditional probability of observing "e" if the previous two letters were "aa" times the probability that "aa" will be the first two letters of a word, plus the conditional probability of "e" if the previous two letters were "ab" times the probability that "ab" will be the first two letters of a word, and so forth.

The code illustrated in Figure 3.3 uses eight code elements. If nothing is known about prior events or about the structure of the code, the probability of observing "5" is

$$p(i=5) = 1/N = 1/8 \qquad [25]$$

However, if the messages were generated according to the rules given in Chapter 3, in the section "Encoding Unequally Distributed Elements," the following conditional probabilities can be calculated: If i is the first element of a string, or if the previous element was "1," then

$$p_c(i=5) = 0 \qquad [26]$$

If the previous element is "2," then

$$p_c(i=5) = 1/4 \qquad [27]$$

Since the value $p_c(i=5)$ is zero for each of the other two conditions, and the previous element is "2" only 1/4 of the time, the overall probability that any randomly chosen elemnt, i, will be 5 is

$$p_c(i=5) = 1/16 \qquad [28]$$

Summed over all three conditions, $p_c(5) = 1/16$. The difference between expected (1/8) and observed frequency (1/16) is due to the structure of the code. Allowing for the effects of structure, the value of H is

$$H_{struct} = -\Sigma p_c(i) \log_2 p_c(i) \qquad [29]$$

summed over all values of $p_c(i)$ for every possible set of conditions, for each element i. The reduction of H as a result of structural constraints is

$$D_{struct} = H_{max} - H_{struct} \quad [30]$$

Structural redundancy can be calculated as

$$R_{struct} = (H_{max} - H_{struct}) / H_{max} \quad [31]$$

$$= 1 - (H_{struct} / H_{max})$$

The constraints imposed by the rules for Code 3′ (see the section "Distributional Redundancy," pp. 38-39) are completely reflected in the distribution of observations, so that $H_{obs} = H_{struct}$ and distributional redundancy is identical to structural redundancy. However, another set of rules might require that an entire message be repeated: Repetition increases structural redundancy to at least 50%, but distributional redundancy might be close to zero. When all of the rules of English are considered, structural redundancy is close to 70%—much higher than the distributional redundancy of 13.7%.

The distinction between distributional and structural redundancy may be illustrated by a code constructed so that every string is composed of 26 letters in alphabetical sequence, beginning with "a." Such a code would have the peculiar characteristics that $H_{obs} = H_{max}$ and $R_{obs} = 0$ while $H_{struct} = 0$ and $R_{struct} = 1$. Although the code set has maximum variety and minimum redundancy in the distribution of elements, in the structure of messages, it has no variety and maximum redundancy.

Note that the foregoing discussion focuses only on the contingent probabilities resulting from internal structure, the rules by which the code assembles its elements into messages. The redundancy produced by external structure—the relationship of code to context—can be calculated by similar procedures if conditional probabilities can be estimated reasonably well. These ideas are taken up in Chapter 5.

*Repetition versus redundancy.* The process of creating a message will frequently produce repetitious sequences that can be easily mistaken for examples of redundancy. For example, in the code of Arabic numerals, the sum of 3010252 and 7090758, 10101010, is repetitious, but it is not redundant. The string 10101010 is exactly as probable as 10238721, or any other sequence of eight

numerals (excluding sequences beginning with "0"). Either numeral may express the net worth of a corporation or the number of bombs dropped by one nation on the citizens of another. The repetition results from the idea that is being expressed, and not from the structure of the code.

Encrypted into a set of elements including "a" and "b," the number of bombs dropped on the citizens of another nation might be transmitted as "babababa," a sequence that could also result from a malfunctioning typewriter or the babbling of an infant. Analysts are sometimes inclined to ascribe a sequence such as "10101010" or "babababa" to redundancy, but there is no basis for doing so without knowing something about the code. Indeed, such repetitious strings probably result more frequently from more or less random events (as in the babbling of an infant or a report of the results of military action) than from the structure of a code. Similarly, the vowel "e" is repeated three times in *elements,* but three different vowels appear in *internal.* Both words are equally good examples of the redundancy produced by the requirement that every syllable in English must contain at least one vowel. (Not all languages have such a requirement. Written Hebrew, for example, uses no vowels at all.) The fact that the first word is also an example of repetition is coincidental and has nothing to do with the fact that it is an example of redundancy. Redundancy always refers to a *general* property of the code; repetition refers to a *particular* property of some message.

# 5. *Structure and Relevance*

Chapter 4 introduced the idea of code structure and distinguished between the internal structure of a code (constraints on the relationship among its elements in a typical message) and its external structure (the relationship of its elements to external events, ideas, and communicative purposes). Chapter 5 will explore the concept of external structure further and link the external structure of a code to the concept of relevance, first introduced in Chapters 1 and 2.

Signal transmission theory explicitly assumes a manufactured signal transmission system. The transmitter and receiver are designed to recognize certain phenomena as signals and ignore the rest. Paul Revere's signaling system established a system in which lights counted as part of the signal if and only if they appeared in the steeple of the Old North Church, a system in which the message could be drawn from only one of two defined elements. Electronic transmission establishes a system in which electromagnetic waves count as part of the signal if and only if they fall within a certain range of frequencies. Human language appears at first glance to satisfy a similar set of assumptions, in that only a few sounds and gestures are recognized; everything else is ignored. At virtually every level, however, language interacts with context and relevance—in short, with meaning.

Relationships among the elements of a code constitute its internal structure; the relationship between elements and meaning constitute the external structure of the code. Between these extremes is the signal transmission system itself. On the one hand, internal structure reflects accommodation to the physical characteristics of a transmission system, such as the signals it can reliably detect, the amount of noise in the system, and its capabilities for encoding and decoding. On the other hand, the signal transmission system itself reflects (by deliberate design or by evolution) a set of accommodations to meaning, to the structure of individual action and social organization.

## The Structure of the Transmission System

The idea of a maximally efficient code ($H_{max}$) presumes that the receiver will be influenced by exactly those events that are part of the code and by nothing else. Consider the telegraphic code consisting of pulses of electricity (dots and dashes or on/off states). The physical design of the telegraphic receiver assures that very close to 100% of the events in the universe will be ignored. The receiver responds only to a small set of events originating in the transmitter. The situation in which a maximally efficient code can even be defined already contains an important quantity of informational structure.

*Establishing structure.* Before any efficient code can be used, a preparatory message must already have been sent: "The code is red versus white beads of a size between two and three centimeters arranged on a wire situated directly between the teller and the other," or "the code is vocalizations that approximate phonemes from a known set," and so on. The selection of medium and code presupposes an extensive prior exchange of information (during the design process), and the communicative event cannot be fully understood without considering this preparatory information. In effect, the advance exchange of information establishes the external structure of a code, its relationship to the signal transmission system and to the meanings to be communicated.

If information has not been exchanged in advance, a more structured message will be required, first to attract attention and then to get some idea across. This is what we do when we use exaggerated gestures to communicate with someone across a large, noisy room. In effect, we are creating a sequence of events with an unusually high degree of internal structure to compensate for a low degree of external structure. Conversely, married couples and intimate friends often communicate effectively by a single word or gesture: Less internal structure is needed because of the high level of external structure in the relationship.

What is most notable about human communication is not the latitude of sounds and gestures we interpret but the much wider latitude of sounds and gestures we ignore. Both internal and external structure contribute to our decisions about what to ignore and what to interpret. A high degree of internal structure leads us to assume that an event has communicative import. When the event is part of a system (e.g., a code) with a known relationship to a context, this external structure leads us to infer communicative import. Early in life we learn that vocalizations and certain gestures are reserved for communication. We minimize noncommunicative uses of these actions and, when others use them, we assume that communication is happening. If a situation is conventionally understood as calling for a communicative response, then any action, even a complete lack of action, will be interpreted as a communicative response: Silence during an interrogation is widely interpreted as a confession of guilt.

*Structure as meaning.* Part of what makes it difficult to apply signal transmission theory to natural communication systems such as language is that natural communication systems routinely incorporate external relationships into the message in a way that obscures the distinctions between structure and content, code and context. The amount of internal structure in a linguistic event is ordinarily reduced in proportion to the external structure available to aid interpretation. In a conversation between two intimates, using the highly structured forms of grammatically correct language would be considered excessively formal or even pompous. However, this very fact can be used to communicate something: Use of a needlessly correct (structurally redundant) turn of phrase is itself a datum. Depending on the context, it may be interpreted as a statement about the quality of the relationship; about the speaker's attitude toward the context, such as if the conversation is taking place in a museum or a pretentious art gallery; or about the possibility that the conversation is being monitored.

The extreme redundancy of language can also be used in such a way that deliberate errors are constructed as data and interpreted as relevant to some context. Members of a minority group may use grammatical "errors" to make fun of the pretensions of a majority. Deliberate falsehoods may be used, in a context in which they are certain to be detected, for purposes of irony or to communicate different ideas to those who are expected to detect the falsehoods as compared with those who are not "in the know." For example, during the 1991 war against Iraq, U.S. and British prisoners of war were interrogated on television. The prisoners used deliberate falsehoods, awkward sentence structure, and expressionless speech patterns to encourage viewers to interpret their "confessions" as the result of coercion.

*External structure and error detection.* The relationship of a message to the context also contributes to error detection and correction and thus reinforces the function of internal structure. For example, consider the cliché, "See ya Monday." If it happens to be the Friday before Memorial Day, the hearer may automatically correct "Monday" to "Tuesday" or, if there is any doubt ask, "You mean Tuesday, right?" Here the structure of language as a code becomes almost inextricably intertwined with the

processes of inference by which events are interpreted, whether they are encoded or not.

*Correcting the estimate of structural redundancy.* If everything that can possibly be known about the context, including the speaker's and the listener's full psychological and social history, could be included in the calculation of conditional probabilities, the structural redundancy of typical messages in a natural language would probably increase well beyond the 70% that has been estimated on the basis of grammatical structure alone. If every behavior is fully caused by antecedent events, as scientists commonly assume, then a full account of external structure would increase structural redundancy to 100% and reduce H to zero. If all such conditional relationships could be included in the calculation, structural redundancy would provide a handy estimate of how well we are able to explain, predict, and control events, and, conversely, H would provide an estimate of how much is left to be explained. Contrarily, the use of codelike features of language for purposes of irony and other noncoded forms of communication has the opposite effect, tending to decrease the effective redundancy and increase H.

*The exclusion of meaning.* Meaning (i.e., external structure) is excluded from signal transmission theory because the effects of external structure on the conditional probabilities of a message are quite difficult to express in a general form that can be usefully quantified and incorporated into the design of a signal transmission system. In the overall context of human communication, the exclusion of external structure is necessarily arbitrary, because the context of a communicative event influences both what is counted as data and how the data are interpreted.

The distinction between data and information is itself one of degree. The less dependence there is on context, the more datalike a message is; the more dependence there is on context, the more the message is like information. Pure information is impossible, and pure data (with no relevance of any sort) cannot be defined. In natural communication systems such as language, the distinction between internal and external structure, code and relevance, is often quite vague. Consider the ritualized, highly predictable phrases used in greeting and parting from acquaintances: Do these belong

to the structure of language as a code, or do they belong to the structure of social relationships as a context of communication?

## The Structure of a Message

Within the constraints of internal structure (the rules by which a code allows messages to be organized), the structure of the message is determined by the external structure of the code and by whatever situation led to the message. The meaning of a message is constrained by the external structure of the code (the relationship of code elements to ideas and events in the context), but the meaning originates in events external to the code and the transmission system proper. The structure of each message combines the influences of the internal and external structure of the code with the structure of a communicative situation in which the message originates and will be interpreted. To understand the structure of any message, the social structure of the context must also be considered.

## Social Structure

It appears that even the most routine forms of human communication can be understood only in the context of the social relationships in which they take place; conversely, those relationships can be observed primarily in the form of communicative interaction. Giddens (1984) expresses the relationship as one in which the structure of social relationships is reproduced (i.e., created and maintained) in the process of routine communicative interaction at the same time that routine communicative interaction is possible only in the context of those same social relationships. It is possible to tell something (convey information) only through the context of common language, expectations, and interpretations. Conversely, when something is told, the context of language, expectations, and interpretations is strengthened or altered, and social structure is given form. In brief, the idea of information as telling something and the idea of information as giving form are, at the social level, different aspects of the same idea.

*Control.* Beniger (1986) expresses the function of communication in terms of control. Control requires perception, comparison, and action: What is the current state of the environment? How does it compare with the preferred state? What action will make the current state more like the preferred state? Societies survive by three processes, which must be kept in balance: production (getting material resources), transportation (moving resources to where they are needed), and consumption (converting resources into energy and structure). When any of these three processes gets out of balance (e.g., consumption outstrips production or transportation can't keep up with production), the result is a *crisis of control,* which can be solved only by developing a new information technology. Because each new technology tends, in time, to strengthen one process more than the others, history advances through a series of control crises, each precipitated by the solution of the previous crisis.

Beniger says, for example, that, in the nineteenth century, a crisis in control grew out of the development of an extensive railroad system. After tracks extended beyond a few miles, the railroad companies were faced with the problem of controlling traffic, as became evident after several spectacular train wrecks. The crisis was resolved through a complex set of innovations, including a greatly expanded bureaucracy (to keep up-to-date records of equipment, shipments, and schedules) and a new communication system, based on a network of telegraph wires connecting various stations (to coordinate the movement of particular trains).

Solution of the transportation crisis followed solution of a previous production crisis by the introduction of steam power and mass production lines; together these two solutions led to a crisis of consumption and a series of economic depressions. The crisis of consumption was resolved through development of the "demand control" technologies of mass advertising, based on advanced print technologies and, later, electronic radio and television broadcasting. The success of demand control has itself precipitated a crisis of information control.

*Preprocessing.* Control requires communication. For example, preventing train collisions requires communication about the location and movement of each train, followed by orders to some trains to wait on sidings while other trains pass them by. Preventing surpluses and scarcities of products such as shoes

and lawn mowers requires communication about the rate of consumption in each market and the rate of production at each source, followed by (a) orders to shift transportation from markets with a surplus to markets with a scarcity, (b) orders to increase or decrease production, and (c) advertising to control the rate of consumption in some or all of the markets.

As control is extended to more aspects of the production, transportation, and consumption system, communication requirements increase, and the burden of processing all the data can rapidly become technically unmanageable. Nass (1988; Beniger & Nass, 1986; Nass & Lee, 1990) has shown that this problem leads to dividing the task into data processing and a preparatory stage of *preprocessing*. Virtually the entire increase in the proportion of "information workers" in the labor force over the last 200 years can be explained by an increase in preprocessing work.

Preprocessing includes anything that is done to make the processing task simpler and easier before data enter a control system. Alphabetizing names is a form of preprocessing. Dividing the nation into time zones was a very powerful form of preprocessing introduced at least in part to facilitate the railroads' control over train schedules. Standardized clothing sizes is a form of preprocessing introduced to facilitate mass production of clothing, and grouping the clothing by style and size on the store shelves is a form of preprocessing that facilitates mass marketing.

*Reducing variety.* Preprocessing works by throwing away information. Time zones throw away the fact that it is actually a few minutes earlier in Lincoln than it is in Omaha. Shoe sizes throw away the infinite gradations in the length, width, thickness, and flexibility of human feet: The customer is either a size 10 or a size 10 1/2, either a C width or a D width. There is nothing between. When preprocessing takes the form of alphabetizing names, imposing standard time zones, and implementing a system of standardized clothing sizes, it reduces the variety in possible messages. According to Beniger, preprocessing also includes bureaucratic regulations restricting the ability of agents to make decisions, regulating the flow of messages, and requiring functionaries such as railroad porters, police, and flight attendants to wear uniforms that clearly indicate their role. Preprocessing of this form reduces the variety in social structure.

Preprocessing strips away all data that are not relevant to a particular task, for example, the task of fitting a human body into a suit of clothing. Preprocessing also constrains the allowable relevance relationships by constraining contexts: The task of fitting a human body into a suit of clothing is further redefined by preprocessing it as the task of fitting a human body into a suit of *mass-produced* clothing. Preprocessing increases the proportion of data to relevance and restructures the contexts by which relevance is determined. By increasing the internal structure and reducing the external structure of the communication system, preprocessing renders the social structure more machinelike, that is, *rationalizes* the relationship of producer to consumer, of worker to manager.

# 6. *Uncertainty*

The relationship between information and H is often discussed in terms of certainty or its opposite, uncertainty (e.g., see Krippendorff, 1975, 1977, 1984; Shannon, 1949, pp. 49-53). Information as a concept in human communication, and H as a quantity in signal transmission, are both said to reduce uncertainty, an idea that is appealing on both intuitive and linguistic grounds. Information is the quality of a message that tells something (Chapters 1 and 2); a remedy for uncertainty about something is to be told about it. If H measures the variety of possible things that might be specified by a message, as well as the power of elements in a message to specify something (Chapter 3), then H must have something to do with the degree to which each element reduces uncertainty about what is specified.

## Defining Uncertainty

*Uncertainty* and *certainty* derive from the Latin *cernere*, "to distinguish, decide, or resolve." Inasmuch as information can be defined in terms of the power of a message to identify or specify

something, and one bit of data is the ability to identify or label one of two subsets within a set of alternatives, it is reasonable to define information as that by which things are ascertained, certainty is increased, and uncertainty (the opposite of certainty) is decreased.

Uncertainty can pertain to (a) an objective description of a particular message, (b) a description of the relationship between a message and a context that is objective in the sense that it is accessible to any alert perceiver, or (c) a subjective description of the relationship between a message and a particular person's cognitions, including what that person already knows. To avoid confusion, it is important to distinguish each of these from the others. It is also important to distinguish between a particular message and the general or "typical" message, between the subjective response of a particular person and an objective description of the array of responses that may be expected of a "typical" person.

Statistical uncertainty has to do with the probability distribution of elements in a set and can be described only for the general observation, logically prior to any particular observation. Subjective uncertainty has to do with what some person expects in a particular situation and can be described either before or after the observation. A particular event may change the subjective uncertainty of any person who observes it but can have no effect on the statistical uncertainty.

*Statistical uncertainty.* Signal transmission theory is concerned primarily with uncertainty in a statistical sense, that is, the dispersion of observed values for some variable over a number of instances, the uncertainty that applies to the typical message. The more elements in the set, the lower the probability of observing any particular element, and the more *statistical* uncertainty is attached to each potential observation. The total statistical uncertainty in a set of elements is simply a different label for H, the measure of dispersion among elements of a set. An element to be drawn from a small set such as {heads tails} will ordinarily have less uncertainty in this statistical sense than an element to be drawn from a larger set such as the roll of a die {1 2 3 4 5 6}. An element from the binary set {heads tails} reduces less statistical uncertainty than an element from the six-element set {1 2 3 4 5 6}.

By way of illustration, imagine that Professor Jones teaches two sections of speech, each with 16 students who sit in four rows of 4 students each. Jones promises to use the following code each day to let the students know who gets to speak. In section A, each student will be identified by a string of four zeroes and ones. The first digit of the string will tell whether the speaker sits in the first or last two rows; the second will tell whether the row number is odd or even; the third digit will tell whether the speaker sits in the first or last two columns; and the fourth digit will tell whether the column is odd or even. In section B, both rows and columns will be numbered 1 to 4; the first digit will identify the row, and the second will identify the column. In section A, four digits are required to identify a particular student. In section B, only two digits are required.

When Jones writes the first digit in section A, the digit represents only one bit and eliminates only half the students from consideration. The second digit eliminates half of those remaining, so the first two digits eliminate a total of three quarters of the possibilities. But in section B, only one digit is required to eliminate three quarters of the students. Because Jones uses a four-element set in section B, each element represents two bits of data and does as much to reduce uncertainty as two elements from the binary set used in section A.

*Statistical versus subjective uncertainty.* Once the entire label has been written on the blackboard, the students no longer have any *subjective* uncertainty as to who will speak first. However, there is no basis for concluding that the information somehow disappears once the students have read the message. The *statistical* uncertainty is unchanged, the data are still on the blackboard, and the data are still relevant to the question of who speaks first: The information is still there. Each element of the four used in section B still has a statistical probability of one fourth, and each element of the two used in section A still has a statistical probability of one half. Probabilities calculated after the fact always have a value of either 1.0 or 0.0 (the event happened or it didn't) and can never be used to calculate values of H. The assertion that a message contains no information once its contents are known (Dretske, 1981; Garner, 1962; Moles, 1958/1966) is a red herring, the result of confusing statistical uncertainty with subjective

uncertainty. Statistical or objective uncertainty always refers to the expected values of possible events prior to a particular observation and must be carefully distinguished from subjective uncertainty.

*Subjective uncertainty.* This refers to unfulfilled expectations or anticipations of future messages. Subjective uncertainty can be aroused by a puzzle such as the card-bead encoding problem (Chapter 4) or a cryptogram and can be alleviated by solving the puzzle. Subjective uncertainty can also take the form of social uncertainty—a condition in which someone is uncertain about what to expect or how to act in a social situation. Uncertainty in the social sense can be aroused by any message or perception that increases the perceived probability of social interaction involving either an unfamiliar person or an unfamiliar context.

## Uncertainty and False Paradoxes

When uncertainty as a subjective description of what a person does or does not already know is confused with H as a statistical description of the distribution of elements in a set, especially when statistical uncertainty prior to an event is confused with subjective uncertainty after the event, the results include false paradoxes and theoretical blind alleys. Sometimes these paradoxes are merely distractions, and lead at worst to a waste of time, but sometimes they discourage people from thinking seriously about information or from using H as a research tool.

*The paradox that noise increases information.* One of the earliest examples of the problems that can arise when H is confused with subjective uncertainty occurs in Weaver's (1949) interpretation of Shannon's theory. Weaver's mistake is worth a brief examination, if only because it is still being repeated in popular writings and even in some communication theory textbooks.

Recall that "noise," in signal transmission theory, refers to any random alteration in the signal. Thus, in a statistical sense, the more noise there is in a channel, the more likely it is that elements in any string will have been randomly altered and the less likely it is that the message selected at one place and time will be

accurately reproduced at another. Noise in the channel increases the observer's *subjective* uncertainty about what message was sent, but it has no effect on H, the dispersion of elements in the code from which the message elements were assembled (the statistical uncertainty). Noise is likely to increase the statistical variety of the signal by equalizing the distribution of probabilities, because the more frequently used elements are more likely to be affected by random processes, but that possibility is not directly related to the information capacity of the code.

The uncertainty associated with the statistical properties of a code describes the expected value of the distribution of elements in a *typical* encoded message. The uncertainty associated with noise in the channel describes the probability that a given element of the received message will be identical to the corresponding element of the transmitted message. Noise in the channel reduces the transmission capacity of the system in two ways: First, it reduces the power of any message to specify something. Second, the only way to combat noise is to increase the structure of the code, thereby increasing redundancy and reducing the variety of the code ($H_{struct}$). The more noise in the channel, the more structural redundancy will be required to achieve a given level of accuracy in transmission and the less data can be transmitted by a message of a given size.

Because Weaver confounded cognitive with statistical uncertainty, the general with the particular, the statistical variety of a code with the statistical variety of a message after it has been transmitted, he concluded that

> when there is noise, the received signal exhibits greater information. . . . This is a situation which beautifully illustrates the semantic trap into which one can fall if he does not remember that "information" is used here with a special meaning that measures freedom of choice and hence uncertainty as to what choice has been made. (Weaver, 1949, p. 19)

In signal transmission theory, H measures the amount of unconstrained variation in a code, hence the average power of each element of the code to specify something. When the definition is extended to measure the capacity of a particular message, it is *still* based on the overall distribution of elements in the code, *not*

on the distribution within the particular message. To the degree that conditions in a transmission channel randomly change elements of a message, they reduce the amount that can be specified by that message and thus reduce the information. Whether the "noise" in the channel increases or decreases the statistical dispersion of elements in the message is irrelevant, because H is calculated on the dispersion of elements in a *typical* message, not in a particular message. It is never the case that information increases as an observer becomes more uncertain about what signal was actually transmitted, and it is certainly never true, as Weaver claimed, that random perturbations in a signal can be "subtracted out" or that they can somehow increase the information content of the signal.

*The paradox of information and meaning.* Weaver's misconstrual of Shannon's theorem concerning the relationship between noise and information, which resulted directly from confounding the two senses of the word *uncertainty* and his leap from the general to the particular, led to a further assertion that "information and meaning may prove to be something like a pair of canonically conjugate variables in quantum theory, they being subject to some joint restriction that condemns a person to the sacrifice of the one as he insists on having much of the other" (Weaver, 1949, p. 28). It has led to the oft-repeated but quite mistaken contention that the *information* in information theory is to be understood not merely in terms of a restricted range of the everyday meaning we assign to the word but as an actual contradiction or opposite of the way the word is used in discussions of human communication (see Ritchie, 1986).

## Managing Uncertainty in Human Communication

The distinctions among H as a measure of variety in a code set, data as a set of particular messages, and information as data relevant to some context provide a basis for understanding the relationship between uncertainty and information seeking. Berger and Calabrese (1975) argue that interpersonal encounters arouse uncertainty and that communication in such encounters can be at least partly explained as a process of reducing uncertainty. Daft

and Weick (1984) make a similar argument about communication activities of organizations and of individuals within organizations.

Uncertainty is not necessarily undesirable; neither is its reduction necessarily a primary objective in human communication. As Berger and Calabrese (1975, p. 101) observe, "while uncertainty reduction may be rewarding up to a point, the ability to completely predict another's behavior might lead to boredom." Similarly, Eisenberg (1984, 1986) has shown that ambiguity can be an important strategic move. In general, unpredictability seems to play a positive role in the development of relationships: Specifying too much too soon can foreclose future options and limit freedom of action. The idea of uncertainty reduction might usefully be included, along with other ideas such as uncertainty increase, in a general conceptualization of communication as the management or control of uncertainty.

*Increasing and decreasing uncertainty.* Subjective uncertainty can be increased by a message that (a) arouses expectations by suggesting a distribution of alternatives ("you will meet a handsome stranger") or (b) changes the relevance of already perceived alternatives ("the person seated at the table in the corner is your dinner companion"). Subjective uncertainty can be decreased by a message that (a) changes the probability structure of alternatives ("the handsome stranger will be a sales agent for Long Life Insurance") or (b) changes the relevance of perceived alternatives ("the person seated at the table in the corner is leaving in three minutes"). Subjective uncertainty is equivalent to statistical uncertainty only in the rare circumstance in which an individual correctly understands the probability distribution of potential events or observations.

Consider a situation in which two students meet at a party, and she asks him, "What is your major?" If the college offers between 100 and 200 distinct majors, his reply statistically represents about seven bits of data. If she is aware of only four equally probable majors, the value of his message for her is two bits. But if all she really wants to know is whether he is materialistically or idealistically inclined, the subjective value of the message is about one bit. From the perspective of her subjective uncertainty, the amount of "information" in his response is dependent on the

way she thinks about the possible development of the relationship: We need to know something about the *relevance* of the response in order to say anything interesting about how it may affect subjective uncertainty.

## Uncertainty in Communication: Causes and Remedies

It appears that there are several different types of uncertainty in any communicative encounter, each of which interacts in a different way with information. Daft and Weick (1984; see also Weick, 1979) propose that organizations engage in interpretation of information to reduce both uncertainty and ambiguity. Much of their argument can also be applied to individuals as information users. They classify information interpreters according to (a) whether they regard the environment as analyzable or unanalyzable and (b) whether they are active or passive in seeking information.

The first dimension corresponds closely to the distinction between a code model and a relevance model. An entity that assumes its environment to be analyzable is subscribing to a code model, with the implicit expectation that, for every question, there is, at least in principle, some corresponding datum. An entity that assumes its environment is not analyzable is subscribing to a relevance-type model, with the implicit expectation that at least some crucial questions may not have unique corresponding answers.

Uncertainty about a message that has been received can arise in at least four ways: (a) The message has been randomly perturbed by noise in the transmission channels; (b) the message is incompletely encoded (nonspecificity); (c) the message has been deliberately constructed to permit more than one interpretation (encoding ambiguity); and (d) the interpreter is uncertain as to the context or the array of alternatives to which a message refers (contextual ambiguity). Table 6.1 summarizes these ideas.

*Transmission errors.* Transmission errors (noise) include any random alterations of message elements. Noise can lead to subjective uncertainty for both the originator and the receiver of a message: The receiver is uncertain about what was intended and

**TABLE 6.1** Uncertainty and Information

| *Cause of Uncertainty* | *Informational Remedy* | |
|---|---|---|
| | *Code Model* | *Relevance Model* |
| Noise; errors in transmission | Internal structure (redundancy) | Internal structure (redundancy) |
| Nonspecificity; incomplete encoding | Complete the message | Complete the message |
| Ambiguous encoding | Add data to specify the relevance | Relevance search within the context |
| Ambiguous context | Add data to specify the context | Activate new contexts |

the originator is uncertain about what was received. The only remedy for noise is to rely on the structural constraints of the code to help identify and correct random alterations in the message. However, it is always possible for random processes to produce a message that appears to conform to the rules of the code. For example, a typist may accidentally substitute a phrase from one letter for a phrase in another. Although the use of a highly redundant code may eliminate the cognitive uncertainty associated with noise altogether, thus producing a false sense of security, the statistical uncertainty can never be entirely eliminated.

*Nonspecificity.* If an employer asks a chronic latecomer, "When did you arrive this morning?" and the offender replies, "I arrived at eight oh nine and thirty three point seven seconds," the reply would be considered socially offensive, even if it was demonstrably accurate. On the other hand, "I arrived at about eight" might be considered evasive and slightly dishonest. It is rare in human communication for any idea to be fully encoded; the result is often that the interpreter of a message is uncertain as to exactly what was intended. Sometimes incomplete encoding occurs because the originator is uncertain what is required: For example, if someone at a social event asks, "When did you arrive?" the response, "About eight," is ordinarily regarded as complete. If the questioner wants to know whether the other person may have observed something that happened right at eight, the response will be regarded as insufficiently encoded, in which

case the solution is simply to request more data: "Before eight or after?" or "Did you arrive in time to see Jim fall into the punch bowl?"

*Strategic ambiguity.* Sometimes messages are deliberately left incomplete or encoded in such a way that two or more interpretations are possible. Eisenberg (1984) suggests that ambiguity is often used as a deliberate strategy. If a friend promises to "meet you at the corner in ten minutes" when there is more than one corner, the ambiguity is probably not deliberate and thus can be considered merely incomplete encoding. When a tavern is named "The Office," the ambiguity is deliberate and provides a ploy for its patrons: "Hi, honey, I'll be at The Office for a couple more hours, then I'll come right home."

Both ambiguity and nonspecificity take the form of incompletely "encoded" messages; the distinction is in the relationship of the originator and/or the interpreter to the message. "About eight" is deliberately ambiguous in response to a supervisor's request for information about when an employee arrived at work; in response to a casual request for information about when someone arrived at a party, it is at most merely incompletely encoded. Whether or not the reply is adequately encoded depends in either case on its relevance to the question, which may not be known to the originator of the response.

Ambiguity may be strategically used, as in the instance of a tavern named "The Office," to encourage a benevolent interpretation of a message that might otherwise lead to conflict. Ambiguity may also be used to minimize the need to reveal personal opinions and preserve the appearance of harmony: "How do you like my new hairstyle?" "It certainly looks different." In bureaucracies, ambiguity may be used to preserve "deniability" for higher-ranking officers by disassociating them from the actions that are required to carry out their orders. "We can't have that toxic waste sitting around the warehouse any longer. Do something with it." Ambiguity is often used when the originator of a message is unsure of the context and consequently unsure how the reply may be interpreted.

In a code model, ambiguity is merely another form of incomplete encoding and is also remedied by obtaining more data, in this case, data that specify the relevance. "I know my hairstyle

looks different; is it different in a good way or a bad way?" "I know you want me to get rid of the toxic waste. Do you want me to haul it to the city dump in the dead of night or spend most of last year's profits disposing of it legally?" In a relevance model, the uncertainty is reduced by a relevance search: The interpreter thinks of all the ways in which the message is likely to be relevant in the current context and chooses the most likely interpretation.

*Ambiguous context.* Another form of ambiguity results when the context is not known or when a message can be interpreted in more than one context (Daft & Lengel, 1986, p. 557; Goffman, 1974). Off-color jokes and puns often employ this sort of ambiguity, and many a spy story turns on a code phrase constructed to have very different meanings according to how the context is understood. In the code model, as with ambiguous relevance, the uncertainty associated with ambiguous context is resolved by seeking data about the intended context. In the relevance model, the uncertainty associated with ambiguous context is resolved by testing the message against various contexts and interpreting it in relation to the context (or contexts) that provide the most likely or satisfying interpretation.

## Measuring Uncertainty

Thus far, H has been described as a measure of the variety among the elements of a set, a function of the statistical uncertainty associated with the probability of observing each element of the set, for example, in a typical message. The statistical uncertainty associated with H has been carefully distinguished from the subjective uncertainty associated with the perception, decoding, and interpretation of a message. Given that H can be used to measure the amount of variance in any set of elements, is there any reason it cannot be used to estimate, analyze, and compare degrees of ambiguity or subjective uncertainty in a communicative situation?

The answer is that H can indeed, at least in principle, be used for such a purpose, subject to a couple of assumptions and a caution. The first assumption is that it must be possible to estimate the array of possible interpretations or contexts as well as the

probabilities that typical research subjects associate with each. That does not mean that it is necessary to know the full range of possibilities: If most are known, it is reasonable to assign some residual value to those that rarely occur. The second assumption is that these possible interpretations or contexts and the subjective probabilities assigned to each are reasonably constant across persons and/or situations. *If* a separate estimate can be obtained for each person in each situation, the second assumption can be discarded.

The caution is that the researcher may confuse assumption with fact and regard the fixed array of possibilities as a feature of the "real world." The risk is that the process of communication will be treated as fixed and mechanical, and the potential for human interpreters to generate and human message originators to evoke entirely novel responses will then be disregarded. Such a mistake is not an inevitable result of applying H to estimating the range of human responses to ambiguity or uncertainty, but it is a trap easy to fall into and important to avoid (Ritchie, 1986).

## *7. Summary: Distinctions and Connections*

Information has to do with the way an act of communication (i.e., a message) tells something or informs someone of something: The concept of information is embedded in the concept of communication. Mechanical communication (e.g., between the encoder and decoder of a telegraph circuit or between two worker ants) can be conceived of as a relatively simple stimulus-response process, and information can be defined in relation to the set of possible stimulus-response pairs. Over the past several decades, it has become evident that human communication is a more complex process. The concept of information, if it is to contribute to developing a theory of human communication, needs to be understood in a way that reflects this complexity.

## Distinctions

Understanding a concept begins with making distinctions. Some of the most important distinctions for understanding the concept of information are summarized in the following paragraphs. Some of these apply to any concept; others are particular to the concept of information.

*Observation, concept, and measurement.* A concept is an abstraction from multiple events, an expression of some common attribute of these events that has explanatory power. A measurement is an expression of the degree to which this attribute is present in one particular event as compared with other events. In the study of communication, observations include multiple instances in which events are paired with responses in some systematic way. Concepts include the concept of a *code,* the *rules* by which the code organizes these communicative events, and the *messages* into which they are organized. Measurements include $H_{max}$, the variety of possible events; $H_{obs}$, the variety of observed events; and $H_{struct}$, the variety of events after the structural constraints imposed by the rules of the code have been taken into account.

The measurement is not the concept: The same statistical formula can be applied to any number of concepts. The concept is not the observation: The same events can support any number of concepts, depending on what attributes are considered. The concept tells the observer what attributes to observe, and the measurement is used by the observer to record or communicate how much of a particular attribute (compared with some standard) was observed in each event.

*The general and the particular.* Theories and concepts connect the general to the particular. A theory expresses a general relationship among attributes common to many particular observations and predicts relationships that will be observed in future particular observations. Each of these attributes is associated with a concept: A concept tells what a particular event has in common with other events of a general type. The applicability of a theory to a particular event is determined by whether the attributes associated with its concepts are observed.

*Code and message, element and datum.* A *code* is a general set of rules for organizing communicative events; a *message* is a particular communicative event, organized according to some code. A code may be the product of deliberate design, in which case its rules are known in advance, or it may be the product of evolutionary processes of one sort or another, in which case it may be possible to infer its rules only by observing many messages. The smallest coherent event used by a code is an *element;* the smallest coherent event in a message is a *datum* (see Note 1 in Chapter 2). Elements are conceptual units in the general description of a code; data are observed units of a particular message. For example, "d" is 1 of 26 elements used by the code of English spelling and "7" is 1 of 10 elements used by the code of Arabic numerals. Each time "d" appears in the message, "Let the dead bury their dead," it is a datum; each time "7" appears in the message "1776," it is a datum.

*Rules, structure, and redundancy.* Regular patterns of relationships among the elements of a code as they appear in typical messages constitute the *structure* of the code. These patterns are generated by the *rules* by which the elements of the code are organized into messages: The rules may be explicit and codified, as in an artificial code, or they may be implicit, as in a natural language. The *internal structure* of a code results from rules that constrain the sequence of elements in a message. For example, the rules of English spelling require at least one vowel in each word and call for "u" to follow "q" (with a small number of exceptions). The *external structure* of a code results from rules that associate elements of the code with external events, such as with elements of other code sets or with meanings. For example, the rules of English spelling associate the sound of a cat with the letters "meow," and the rules of English vocabulary associate the idea of a feathered biped with "bird."

The structure of a code can be described statistically in terms of conditional relationships between one kind of event and other events. Internal structure can be entirely described in terms of the conditional probabilities associated with each element of the code set, given that certain elements have already been observed in a message. External structure can be described in terms of the conditional probabilities associated with each element (or message),

given that certain events have been observed in the context in which a communicative event takes place.

Just as the structure of a code can be described in terms of the organization of elements in a typical message, the structure of a particular message can also be described in terms of the organization of its data. The structure of a particular message is influenced by the internal structure of the code, by the external structure of the code, and by the structure of the context of the message—that is, by what the message is intended to communicate.

The internal structure of a code restricts the way the elements of the code set can be combined into messages: It makes some combinations more likely, some combinations less likely, and totally eliminates many combinations. For example, *the* is the most common word in the English language, *wan* is uncommon, and *zrd* is not allowed. The resultant reduction in variety, and consequently in data capacity, can be measured by the redundancy of the code.

The distinction between rules and structure is probably less important than the distinction between structure and redundancy. Both rules and structure are conceptual, and they differ at the level of theory. Redundancy is statistical, and the distinction between structure and redundancy is an example of the distinction between concept and measurement. The difference between the structure of a code and the structure of a particular message is also quite important. Redundancy can be computed from the structure of the code but *not* from the structure of any particular message. Thus redundancy must also be distinguished from repetition. Redundancy is a general attribute of a code and repetition is a particular attribute of a message.

*Statistical and subjective uncertainty.* H measures the variety of a code set and the data capacity of a code. Redundancy measures the degree to which code structure reduces data capacity. H and redundancy are both expressed in terms of the probability distribution of code elements in a typical message, that is, the *statistical uncertainty* associated with the probability of observing each element, given the other elements that have already been observed in a message. *Subjective uncertainty* refers to an individual's expectation about an event that has not yet been observed, such as a message that has yet to be received or the

meaning of a message that has been received but has not been understood.

Statistical uncertainty is calculated only for the general case and is logically prior to the observation of any particular event (after an observation, the probability is always equal to either 1 or 0). Subjective uncertainty refers to the particular case. Subjective uncertainty matches statistical uncertainty only in the rare instance in which an individual correctly understands the distribution of probabilities associated with a set of possible events. Statistical measures such as H and redundancy refer only to statistical uncertainty; any connection of such measures to subjective uncertainty is indirect.

*Data and information.* Each time an element of a code appears in a message, it constitutes a datum. Each letter in a typed sentence is a datum with respect to spelling; each word is a datum with respect to vocabulary. To the extent that it can be represented as an element in some set, the sentence itself may be a datum. More generally, any patterns of objects or behavior arranged so as to call attention to communicative potential constitute data, whether or not they belong to a recognized set. Data refers, generally, to communicative events without regard to their meaning. As these events are interpreted and understood, relevant to the context of some human purposes, they constitute *information.*

In a mechanistic system such as transmissions between the computer keyboard and screen, there is no interpretation; the distinction between data and information means nothing. In human communication, the identical message can tell different people very different things: The *relevance* of the data in the message is crucial to understanding information in human communication. "Fire" is a datum. It assumes very different relevance, and is information of a very different sort, in the context of a military battle or a crowded theater.

## Connections

The function of concept explication is to make distinctions, but the purpose of theory construction is to make connections. This volume has emphasized a set of conceptual distinctions based on

an implicit theory that human communication is more than a set of highly complex stimulus-response pairings, that human communication involves interpreting messages according to their relevance in the context of human purposes. This implicit theory suggests a useful metaphor for science, based on the idea that science is a process of making observations according to certain rules and interpreting these data in some theoretical context.

A theoretical perspective is a set of rules for connecting observations with concepts and for combining concepts into theoretical propositions. An observation is a datum, definable within the set of events that might have been observed, and becomes information insofar as it is relevant to a conceptual scheme. A concept is itself informative only within the context of some theory. The concepts discussed in this volume, such as "information" and "uncertainty," are *informative* only by way of their relevance to the context of a theory about human communication.

*How distinctions improve connections.* The exercise of drawing conceptual distinctions contributes to the theory-building process of making connections in several ways. In the first place, careful distinctions among the various attributes of a class of events improve the observer's ability to make predictions about which attributes really matter. Equally important, at least from the perspective of the empiricist, these predictions can *only* be tested if the concepts of a theory are specific enough to discriminate between events that do and events that do not satisfy the assumptions of the theory. Shannon's (1956) warning about the dangers of confusing the "information" of signal transmission theory with the "information" of human communication is still valid. Distinguishing between information and data provides a basis for studying the independent contributions of interpretation, context, and relevance to telling and informing and also provides a firmer foundation for studying the degree to which the data capacity of a message constrains the process of communication.

Distinguishing H from information also broadens the range of possible applications. H can be applied to situations that have little if anything to do with information. For example, H can be used to measure the variety in public affairs agendas of communities (Chaffee & Wilson, 1977). H may also prove applicable in ways that will contribute indirectly to our understanding of

information. For example, it may prove useful to estimate the variety of contexts in which a particular message is commonly considered relevant within some community or the variety of messages that may be used in some context. It may also prove useful to obtain estimates of the variety *perceived* by subjects (i.e., what is the perceived range of contexts or responses and what is the perceived probability that "a typical person might respond in that way") and to explain individual communicative behavior accordingly. How to estimate the variety of some set is a methodological question. Which if any of these instances of variety might be worth measuring is a theoretical question. To assume that H and various theorems using H must be useful in human communication research because they are useful in the theory of electronic circuits is to carry an interesting metaphor much too far.

The concept of information can be defined and used in theoretical contexts in which H would be awkward or unmanageable. Once information is understood as a function of the relevance of data to some context, the researcher can feel confident in defining and using the concept of information with no need to say anything about variety, H, entropy, or uncertainty: These other concepts can be explained if they are important to the theory under development and ignored if they are not. The only question that cannot be avoided in any discussion of information is the basic question: How does the event inform?

# *References*


Abbott, E. W. (1989). The electronic farmer's marketplace: New technologies and agricultural information. *Journal of Communication, 39*(3), 124-136.

Arbib, M. A., & Hesse, M. B. (1986). *The construction of reality.* Cambridge: Cambridge University Press.

Baxter, L. A., & Wilmot, W. W. (1984). "Secret tests": Social strategies for acquiring information about the state of the relationship. *Human Communication Research, 11,* 171-201.

Beniger, J. R. (1986). *The control revolution: Technological and economic origins of the information society.* Cambridge, MA: Harvard University Press.

Beniger, J., & Nass, C. I. (1986). Preprocessing: Neglected component of sociocybernetic models. In R. F. Geyer & J. van der Zouwen (Eds.), *Sociocybernetic paradoxes* (pp. 119-130). London: Sage.

Berger, C. R. (1979). Beyond initial interaction: Uncertainty, understanding, and the development of interpersonal relationships. In H. Giles & N. St. Clair (Eds.), *Language and social psychology* (pp. 122-144). Oxford: Basil Blackwell.

Berger, C. R., & Calabrese, R. J. (1975). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human Communication Research, 1,* 99-112.

Berger, C. R., Gardner, R. R., Parks, M. R., Schulman, L., & Miller, G. R. (1976). Interpersonal epistemology and interpersonal communication. In G. R. Miller (Ed.), *Explorations in interpersonal communication* (pp. 149-171). Beverly Hills, CA: Sage.

Bittner, J. R. (1988). *Fundamentals of communication* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Borman, E. G. (1989). *Communication theory.* Salem, WI: Sheffield.

Braman, S. (1989). Information and socioeconomic class in U.S. constitutional law. *Journal of Communication, 39*(3), 163-179.

Bransford, J. D., & McCarrell, N. S. (1977). A sketch of a cognitive approach to comprehension: Some thoughts about what it means to comprehend. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking: Readings in cognitive science* (pp. 377-399). Cambridge: Cambridge University Press.

Brooks, W. D., & Heath, R. W. (1989). *Speech communication* (6th ed.). Dubuque, IA: William C Brown.

Busby, R. E., & Majors, R. E. (1987). *Basic speech communication: Principles and practices.* New York: Harper & Row.

Chaffee, S. H., & Wilson, D. G. (1977). Media rich, media poor: Two studies of diversity in agenda-holding. *Journalism Quarterly, 54,* 466-476.

Cherry, C. (1965). The communication of information. In A. G. Smith (Ed.), *Communication and culture* (pp. 35-40). New York: Holt, Rinehart & Winston.

Cherry, C. (1982). *On human communication: A review, a survey, and a criticism* (3rd ed.). Cambridge: MIT Press. (Original work published 1957)

Couch, C. (1988). Mathematics as information technologies. *Journal of Communication, 32*(2), 33-48.
Daft, R. D., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science, 32*, 207-224.
Daft, R. L., & Weick, K. E. (1984). Toward a model of organizations as interpretation systems. *Academy of Management Review, 9*, 284-295.
Darnell, D. K. (1970). Clozentropy: A procedure for testing English language proficiency of foreign students. *Speech Monographs, 37*, 36-46.
DeFleur, M. L., & Ball-Rokeach, S. (1989). *Theories of mass communication* (5th ed.). New York: Longman.
DeFleur, M. L., & Dennis, E. E. (1988). *Understanding mass communication* (3rd ed.). Boston: Houghton-Mifflin.
DeVito, J. A. (1988). *Human communication: The basic course* (4th ed.). New York: Harper & Row.
Dewart, L. (1989). *Evolution and consciousness: The role of speech in the origin and development of human nature.* Toronto: University of Toronto Press.
Dewey, J. (1948). *Reconstruction in philosophy.* Boston: Beacon. (Original work published 1921)
Dickens, M., & Williams, F. (1964). An experimental application of "cloze" procedure and attitude measures to listening comprehension. *Speech Monographs, 31*, 103-108.
Douglas, W. (1985). Anticipated interaction and information seeking. *Human Communication Research, 12*, 243-258.
Dretske, F. I. (1981). *Knowledge and the flow of information.* Cambridge: MIT Press.
Eisenberg, E. M. (1984). Ambiguity as strategy in organizational communication. *Communication Monographs, 51*, 227-242.
Eisenberg, E. M. (1986). Meaning and interpretation in organizations. *Quarterly Journal of Speech, 72*, 88-97.
Feldman, M. S., & March, J. G. (1981). Information in organizations as signal and symbol. *Administrative Science Quarterly, 26*, 171-186.
Finn, S. (1986). Unpredictability as correlate of reader enjoyment of news articles. *Journalism Quarterly, 62*, 334-339.
Finn, S., & Roberts, D. F. (1984). Source, destination, and entropy: Reassessing the role of information theory in communication research. *Communication Research, 11*, 453-476.
Frick, F. C. (1959). Information theory. In S. Koch (Ed.), *Psychology: A study of a science* (Vol. 2, pp. 611-636). New York: McGraw-Hill.
Gandy, O. H., Jr. (1989). The surveillance society: Information technology and bureaucratic social control. *Journal of Communication, 39*(3), 61-76.
Gardner, H. (1987). *The mind's new science: A history of the cognitive revolution.* New York: Basic Books.
Garner, W. R. (1962). *Uncertainty and structure as psychological concepts.* New York: John Wiley.
Garner, W. R. (1974). *The processing of information and structure.* Potomac, MD: Lawrence Erlbaum.
Giddens, A. (1984). *The constitution of society.* Berkeley: University of California Press.
Goffman, E. (1974). *Frame analysis: An essay on the organization of experience.* New York: Harper & Row.
Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics: 3. Speech acts* (pp. 41- 58). New York: Academic Press.
Griffith, B. C. (1989). Understanding science: Studies of communication and information. *Communication Research, 16*, 600-614.

Hartley, R. V. L. (1928). Transmission of information. *Bell System Technical Journal, 7,* 535.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: John Wiley.

Hewes, D. E., Graham, M. L., Monsour, M., & Doelger, J. A. (1989). Cognition and social information-gathering strategies: Reinterpreting assessment in second-guessing. *Human Communication Research, 16,* 297-321.

Hsia, H. (1968a). Effects of noise and difficulty level of input information in auditory, visual, and audiovisual information processing. *Perceptual and Motor Skills, 26,* 99-105.

Hsia, H. (1968b). On channel effectiveness. *AV Communication Review, 16,* 245-267.

Hsia, H. (1968c). Output, error, equivocation, and recalled information in auditory, visual, and audiovisual information processing with constraint and noise. *Journal of Communication, 18,* 325-353.

Hsia, H. (1969). Intelligence in auditory, visual, and audiovisual information processing. *AV Communication Review, 17,* 272-282.

Hybels, S., & Weaver, R. L., II. (1986). *Communicating effectively.* New York: Random House.

Jester, R. E. (1968). Comments on Hsia's auditory, visual, and audiovisual information processing. *Journal of Communication, 18,* 346-349.

Kellermann, K. (1986). Anticipation of future interaction and information exchange in initial interaction. *Human Communication Research, 13,* 41-75.

Krippendorff, K. (1975). Information theory. In G. J. Hanneman & W. J. McEwen (Eds.), *Communication and behavior* (pp. 351-389). Reading, MA: Addison-Wesley.

Krippendorff, K. (1977). Information systems: An overview. In B. D. Ruben (Ed.), *Communication yearbook* (Vol. 1, pp. 149-171). New Brunswick, NJ: Transaction.

Krippendorff, K. (1984). Paradox and information. In B. Dervin & M. Voigt (Eds.), *Progress in communication sciences* (Vol. 5, pp. 45-72). Norwood, NJ: Ablex.

Krippendorff, K. (1986). *Information theory: Structural models for qualitative data.* Beverly Hills, CA: Sage.

Krull, R. J., Watt, J. H., & Lichty, L. W. (1977). Entropy and structure: Two measures of complexity in television programs. *Communication Research, 4,* 61-86.

Kurke, L. B., Weick, K. E., & Ravlin, E. C. (1989). Can information loss be reversed? Evidence for serial reconstruction. *Communication Research, 16,* 3-24.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by.* Chicago: University of Chicago Press.

Littlejohn, S. W. (1989). *Theories of human communication* (3rd ed.). Belmont, CA: Wadsworth.

Lowry, D. T., & Marr, J. T. (1975). Clozentropy as a measure of international communication comprehension. *Public Opinion Quarterly, 39,* 301-312.

Lynch, F. D. (1974). Clozentropy: A new technique for analyzing audience response to film. *Speech Monographs, 41,* 245-252.

Machlup, F., & Mansfield, U. (Eds.). (1983). *The study of information: Interdisciplinary messages.* New York: John Wiley.

Moles, A. (1966). *Information theory and esthetic perception* (J. E. Cohen, Trans.). Chicago: University of Illinois Press. (Original work published 1958)

Murdock, G., & Golding, P. (1989). Information poverty and political inequality: Citizenship in the age of privatized communications. *Journal of Communication, 39*(3), 180-195.

Nass, C. I. (1988). Work, information, and information work: A retrospective and prospective framework. In I. H. Simpson & R. L. Simpson (Eds.), *Research in the sociology of work: Vol. 4. High-tech work* (pp. 311-333). Greenwich, CT: JAI.

Nass, C. I., & Lee, P. (1990). *Information work in the United States, 1900-1980: An information processing model.* Unpublished paper, Stanford University.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning.* Urbana: University of Illinois Press.

Pagels, H. (1988). *The dreams of reason: The computer and the rise of the science of complexity.* New York: Simon & Schuster.

Paisley, W. J. (1966). The effects of authorship, topic, structure, and time of composition on letter redundancy in English texts. *Journal of Verbal Learning and Verbal Behavior, 5,* 28-34.

Parks, M. R., & Adelman, M. B. (1983). Communication networks and the development of romantic relationships: An expansion of uncertainty reduction theory. *Human Communication Research, 10,* 55-80.

Pearson, J. C., & Nelson, P. E. (1988). *Understanding and sharing: An introduction to speech communication* (4th ed.). Dubuque, IA: William C Brown.

Peters, J. D. (1984). *From hylemorphism to hubris: A history and analysis of "information."* Paper presented at the International Communication Association Annual Meeting, San Francisco.

Peters, J. D. (1986). Institutional sources of intellectual poverty in communication research. *Communication Research, 13,* 527-559.

Piatelli-Palmarini, M. (1980). *Language and learning: The debate between Jean Piaget and Noam Chomsky.* Cambridge, MA: Harvard University Press.

Pierce, J. R., & Noll, A. M. (1990). *Signals: The science of telecommunications.* New York: Scientific American Library.

Planalp, S., & Honeycutt, J. M. (1985). Events that increase uncertainty in personal relationships. *Human Communication Research, 11,* 593-604.

Premack, D. (1976). Mechanisms of intelligence: Preconditions for language. In S. R. Harnad et al. (Eds.), *Origins and evolution of language and speech* (pp. 544-561). New York Academy of Sciences.

Richards, I. A. (1936). *The philosophy of rhetoric.* New York: Oxford University Press.

Ritchie, L. D. (1986). Shannon and Weaver: Unravelling the paradox of information. *Communication Research, 13,* 278-298.

Ritchie, L. D. (1991). Another turn of the information revolution: Data, relevance, and the information society. *Communication Research, 18,* 412-427.

Schiller, D. (1989). Informational bypass: Research library access to U.S. telecommunications periodicals. *Journal of Communication, 39*(3), 104-109.

Schramm, W. (1955). Information theory and mass communication. *Journalism Quarterly, 32,* 131-146.

Schramm, W. (1971). The nature of communication between humans. In W. Schramm & D. F. Roberts (Eds.), *The process and effects of mass communication* (pp. 1-53). Urbana: University of Illinois Press.

Schramm, W. (1973). *Men, messages, and media: A look at human communication.* New York: Harper & Row.

Seiler, W. J. (1988). *Introduction to speech communication.* Glenview, IL: Scott, Foresman.

Severin, W. J., & Tankard, J. W. (1988). *Communication theories* (2nd ed.). New York: Longman.

Shannon, C. (1938). A symbolic analysis of relay and switching circuits. *Transactions of the American Institute of Electrical Engineers, 57,* 1-11.

Shannon, C. (1948, July & October). The mathematical theory of communication. *Bell System Technical Journal,* pp. 379-423, 623-656.

Shannon, C. (1949). The mathematical theory of communication. In C. Shannon & W. Weaver, *The mathematical theory of communication* (pp. 31-125). Urbana: University of Illinois Press.

Shannon, C. (1956). The bandwagon. *IRE Transactions on Information Theory, 2,* 3.

Shannon, C., & Weaver, W. (1949). *The mathematical theory of communication.* Urbana: University of Illinois Press.

Slater, M. (1990). Processing social information in messages: Social group familiarity, fiction versus nonfiction, and subsequent beliefs. *Communication Research, 17,* 327-343.

Smith, A. G. (1966). *Communication and culture.* New York: Holt, Rinehart & Winston.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition.* Cambridge, MA: Harvard University Press.

Sperber, D., & Wilson, D. (1987). Precis of relevance: Communication and cognition, with discussion following. *Behavioral and Brain Sciences, 10,* 697-754.

Stubbs, M. (1980). *Language and literacy: The sociolinguistics of reading and writing.* London: Routledge & Kegan Paul.

Taylor, W. L. (1953). "Cloze procedure": A new tool for measuring readability. *Journalism Quarterly, 30,* 415-433.

Taylor, W. L. (1956). Recent developments in the use of cloze procedure. *Journalism Quarterly, 33,* 42-48.

Terwilliger, R. F. (1968). *Meaning and mind: A study in the psychology of language.* New York: Oxford University Press.

Tubbs, S. L., & Moss, J. (1983). *Human communication* (4th ed.). New York: Random House.

Watt, J. H., & Krull, R. (1974). An information theory measure for television programming. *Communication Research, 1,* 44-64.

Watt, J. H., & Welch, A. J. (1983). Effects of static and dynamic complexity on children's attention and recall of television instruction. In J. Bryant & D. R. Anderson (Eds.), *Children's understanding of television.* New York: Academic Press.

Weaver, W. (1949). Recent contributions to the mathematical theory of communication. In C. Shannon & W. Weaver, *The mathematical theory of communication* (pp. 1-28). Urbana: University of Illinois Press.

Weick, K. E. (1979). *The social psychology of organizing.* Reading, MA: Addison-Wesley.

Wiener, N. (1950). *The human use of human beings: Cybernetics and society.* New York: Houghton-Mifflin.

Wilson, E. O. (1971). *The insect societies.* Cambridge, MA: Harvard University Press.

# *Author Index*


Abbott, E. W., 4, 68
Adelman, M. B., 71
Arbib, M. A., 68

Ball-Rokeach, S., 3, 69
Baxter, L. A., 3, 68
Beniger, J. R., 4, 48, 49, 68
Berger, C. R., 55, 56, 68
Bittner, J. R., 4, 68
Borman, E. G., 4, 68
Braman, S., 4, 68
Bransford, J. D., 68
Brooks, W. D., 4, 68
Busby, R. E., 3, 68

Calabrese, R. J., 55, 56, 68
Chaffee, S. H., 8, 66, 68
Cherry, C., 4, 68
Couch, C., 4, 68

Daft, R. D., 60, 68
Daft, R. L., 55, 57, 69
Darnell, D. K., 8, 69
DeFleur, M. L., 3, 4, 69
Dennis, E. E., 3, 4, 69
DeVito, J. A., 3, 69
Dewart, L., 69
Dewey, J., 7, 69
Dickens, M., 8, 69
Doelger, J. A., 3, 70
Douglas, W., 3, 69
Dretske, F. I., 52, 69

Eisenberg, E. M., 56, 59, 69

Feldman, M. S., 69
Finn, S., 8, 69
Frick, F. C., 69

Gandy, Oscar H., Jr., 4, 69
Gardner, H., 6, 69
Gardner, R. R., 68
Garner, W. R., 8, 52, 69
Giddens, A., 47, 69
Goffman, E., 60, 69
Golding, P., 4, 70
Graham, M. L., 3, 70
Grice, H. P., 16, 69
Griffith, B. C., 4, 69

Hartley, R. V. L., 5, 69
Heath, R. W., 4, 68
Heider, F., 70
Hesse, M. B., 68
Hewes, D. E., 3, 70
Honeycutt, J. M., 71
Hsia, H., 8, 70
Hybels, S., 3, 70

Jester, R. E., 8, 70
Johnson, M., 9, 16, 70

Kellermann, K., 3, 70
Krippendorff, K., 32, 50, 70
Krull, R. J., 8, 70, 72
Kurke, L. B., 4, 70

Lakoff, g., 9, 16, 70

Lee, P., 49, 71
Lengel, R. H., 60, 68
Lichty, L. W., 8, 70
Littlejohn, S. W., 4, 70
Lowry, D. T., 8, 70
Lynch, F. D., 8, 70

Machlup, F., 70
Majors, R. E., 3, 68
Mansfield, U., 70
March, J. G., 69
Marr, J. T., 8, 70
McCarrell, N. S., 68
Miller, G. R., 68
Moles, A., 52, 70
Monsour, M., 3, 70
Moss, J., 3, 72
Murdock, G., 4, 70

Nass, C. I., 49, 71
Nelson, P. E., 3, 71
Noll, A. M., 8, 71

Osgood, C. E., 71

Pagels, H., 37, 71
Paisley, W. J., 8, 71
Parks, M. R., 68, 71
Pearson, J. C., 3, 71
Peters, J. D., 71
Piatelli-Palmarini, M., 71
Pierce, J. R., 8, 71
Planalp, S., 71
Premack, D., 71

Ravlin, E. C., 4, 70
Richards, I. A., 16, 71
Ritchie, L. D., 55, 61, 71
Roberts, D. F., 8, 69

Schiller, D., 4, 71
Schramm, W., 4, 71
Schulman, L., 68
Seiler, W. J., 3, 71
Severin, W. J., 4, 71
Shannon, C., 4, 5, 6, 7, 8, 12, 13, 15, 36, 50, 53, 66, 71, 72
Slater, M., 3, 72
Smith, A. G., 8, 72
Sperber, D., 16, 17, 72
Stubbs, M., 38, 72
Suci, G. J., 71

Tankard, J. W., 4, 71
Tannenbaum, P. H., 71
Taylor, W. L., 8, 72
Terwilliger, R. F., 16, 72
Tubbs, S. L., 3, 72

Watt, J. H., 8, 70, 72
Weaver, R. L., II, 3, 70
Weaver, W., 4, 7, 53, 54, 55, 72
Weick, K. E., 4, 56, 57, 69, 70
Welch, A. J., 8, 72
Wiener, N., 7, 72
Williams, F., 8, 69
Wilmot, W. W., 3, 68
Wilson, D. G., 8, 16, 17, 66, 68, 72
Wilson, E. O., 18, 72

# *About the Author*

**L. David Ritchie** (Ph.D., Stanford University, 1987) is Assistant Professor in the Department of Speech Communication at Portland State University, Portland, Oregon. He has also taught at the University of Wisconsin—Madison and at the University of Puget Sound. His research interests include information and society, family communication, and media effects. He has published articles in *Communication Research* and *Journal of Communication*. He has coedited a special issue of *Communication Research* on "micro-macro issues in communication research" devoted to articles focusing on research and theory building across levels of analysis.

# NOTES

# NOTES

# NOTES

# NOTES

# NOTES

Printed in the United States
1535